

QY 901 QEDRSSGTTTELCHVTDERNALRRSSAAHTSHNTYNFTKSENSNRCTSCMPYAKLEYKRSS 960  
 Db 900 QEDRSSGTTTELCHVTDERNALRRSSAAHTSHNTYNFTKSENSNRCTSCMPYAKLEYKRSS 959  
 QY 961 NDSLNSVSDCYGRGQWKPSIESYSEDDSKFCSCYGOYPADLAHKTHSANHWDDNDGE 1020  
 Db 960 NDSLNSVSDCYGRGQWKPSIESYSEDDSKFCSCYGOYPADLAHKTHSANHWDDNDGE 1019  
 QY 1021 LPTPINYSILKYDEQLNSGRQSPSQRWARPKHIIIEDEIKQSEQRQRNQSSTTYPVYTE 1080  
 Db 1020 LPTPINYSILKYDEQLNSGRQSPSQRWARPKHIIIEDEIKQSEQRQRNQSSTTYPVYTE 1079  
 QY 1081 STDDKHLKFPQHFQGOECVSPYRSRGANGSETNRVGSNHGINONVQSLOEDDDYEDDKP 1140  
 Db 1080 STDDKHLKFPQHFQGOECVSPYRSRGANGSETNRVGSNHGINONVQSLOEDDDYEDDKP 1139  
 QY 1141 TNSYERYSEEBEHEERTNYSIKYNEEKHVDQPIDYSLKYATDIPSSQKQSPSPFSKS 1200  
 Db 1140 TNSYERYSEEBEHEERTNYSIKYNEEKHVDQPIDYSLKYATDIPSSQKQSPSPFSKS 1199  
 QY 1201 SSGQSKTEHMSSENTSTPSGNAKRQNLHPSSAQSRSGOPQKAATCKYSSINQETIQ 1260  
 Db 1200 SSGQSKTEHMSSENTSTPSGNAKRQNLHPSSAQSRSGOPQKAATCKYSSINQETIQ 1259  
 QY 1261 TYCVEDTPICFGRCSLSLSAEDEIGCNQTTQADGANTLQIAEIKGIGTRSAEDPV 1320  
 Db 1260 TYCVEDTPICFGRCSLSLSAEDEIGCNQTTQADGANTLQIAEIKGIGTRSAEDPV 1319  
 QY 1321 SEVPASVQHPRTKSSRLQSSLSSESAHKAVEFPSSGAKSPSKGAQTPKSPPEHYVQET 1380  
 Db 1320 SEVPASVQHPRTKSSRLQSSLSSESAHKAVEFPSSGAKSPSKGAQTPKSPPEHYVQET 1379  
 QY 1381 PLMFBSCTSVSLDSFESISASSVQSEPCSGMVSGIISPSLDPSQGTWPSPSKTTP 1440  
 Db 1380 PLMFBSCTSVSLDSFESISASSVQSEPCSGMVSGIISPSLDPSQGTWPSPSKTTP 1439  
 QY 1441 PPTQATQREVPKNAKAPTAERKESGPKQAAVNAAVQVQLPDAATLHFAFESTPDGF 1500  
 Db 1440 PPTQATQREVPKNAKAPTAERKESGPKQAAVNAAVQVQLPDAATLHFAFESTPDGF 1499  
 QY 1501 SCSSLSALSALDBPPTOKDVEIRIMPPVOENDNGNETESEQKESNENOEKAEKTIIDSE 1560  
 Db 1500 SCSSLSALSALDBPPTOKDVEIRIMPPVOENDNGNETESEQKESNENOEKAEKTIIDSE 1559  
 QY 1561 KDLDDSDDDDDIIELEECIISAMPTKSRGKKPAQATASKLPPPVARKPSQLPVVKLLPS 1620  
 Db 1560 KDLDDSDDDDDIIELEECIISAMPTKSRGKKPAQATASKLPPPVARKPSQLPVVKLLPS 1619  
 QY 1621 QNRLOPQKHVSFTPGDDMPRVYCEGTPINFSTATSLSDLTIESPNELAAEGVGRGGAQ 1680  
 Db 1620 QNRLOPQKHVSFTPGDDMPRVYCEGTPINFSTATSLSDLTIESPNELAAEGVGRGGAQ 1679  
 QY 1681 SGFEFKEDTIPTEGRSTDEAOGKTSVTPILDDNKABEGDILAEICINSAMPKGSHPK 1740  
 Db 1680 SGFEFKEDTIPTEGRSTDEAOGKTSVTPILDDNKABEGDILAEICINSAMPKGSHPK 1739  
 QY 1741 FRVKKINDQVQOASASSAPNKNQLDGKKKPTSPVKPIQNTFYRTRVRKNADSNNLN 1800  
 Db 1740 FRVKKINDQVQOASASSAPNKNQLDGKKKPTSPVKPIQNTFYRTRVRKNADSNNLN 1799  
 QY 1801 AERFVSDNKSQKQNLKQNSKDFNDKLPNNEDVRGSPAFDSPHHTPIEGTPCYFSRND 1860  
 Db 1800 AERFVSDNKSQKQNLKQNSKDFNDKLPNNEDVRGSPAFDSPHHTPIEGTPCYFSRND 1859  
 QY 1861 SLSLDFDDDDVDLSREKALRKAKENKESAEKVTSHTELTSNQOQANKTQAIKQPINR 1920  
 Db 1860 SLSLDFDDDDVDLSREKALRKAKENKESAEKVTSHTELTSNQOQANKTQAIKQPINR 1919  
 QY 1921 GQPKPILQKQSTFPQSSKIDPRGAATDEKLQNFATIENTPVCFSHNSSLSLSDIOQENN 1980  
 Db 1920 GQPKPILQKQSTFPQSSKIDPRGAATDEKLQNFATIENTPVCFSHNSSLSLSDIOQENN 1979

QY 1981 NKENEPKETEPPDPSQGBSPKQASGYAPKSPHEVEDTPVCFGRNSSLSLSIDSEDDLLQ 2040  
 Db 1980 NKENEPKETEPPDPSQGBSPKQASGYAPKSPHEVEDTPVCFGRNSSLSLSIDSEDDLLQ 2039  
 QY 2041 ECISAMPKPKKSRLLKGNEXHSPRNMGGIILGEDLTLKDITORPDSHGLSPDSENF 2100  
 Db 2040 ECISAMPKPKKSRLLKGNEXHSPRNMGGIILGEDLTLKDITORPDSHGLSPDSENF 2099  
 QY 2101 WKAIQEGANSIVSLHQAAAAACLSRQASDSDSIISLSGSIISGSPFHLTPQEEKPFT 2160  
 Db 2100 WKAIQEGANSIVSLHQAAAAACLSRQASDSDSIISLSGSIISGSPFHLTPQEEKPFT 2159  
 QY 2161 SNKGPRILKPEGEKSTLTETKIIESKGIKGGKVKYKSLITGKVRNSSEISGQKQPLQAN 2220  
 Db 2160 SNKGPRILKPEGEKSTLTETKIIESKGIKGGKVKYKSLITGKVRNSSEISGQKQPLQAN 2219  
 QY 2221 MPISIRGRTHIHPGVNNSSTSPVSKGPPKPTKPAKSPSEGOATATSPRGAKS VKS 2280  
 Db 2220 MPISIRGRTHIHPGVNNSSTSPVSKGPPKPTKPAKSPSEGOATATSPRGAKS VKS 2279  
 QY 2281 ELSVARTSOIGSSKAPSRSGSRDSTPGRPAQQLSRPIQSPGRNISPGRNGISPPN 2340  
 Db 2280 ELSVARTSOIGSSKAPSRSGSRDSTPGRPAQQLSRPIQSPGRNISPGRNGISPPN 2339  
 QY 2341 KLSQLPRTSSPSTASTKSSGSGKMSYTPSGRQMSQQLTKQTGLSKNASSIPRSESASKG 2400  
 Db 2340 KLSQLPRTSSPSTASTKSSGSGKMSYTPSGRQMSQQLTKQTGLSKNASSIPRSESASKG 2399  
 QY 2401 LNQMNGNGANKKVELSRMSTKSSGSDRSERPVLVROSTPIKEAPSTPLRKLSESA 2460  
 Db 2400 LNQMNGNGANKKVELSRMSTKSSGSDRSERPVLVROSTPIKEAPSTPLRKLSESA 2459  
 QY 2461 SFESLSPSSRPASPTRQAOQTPVLSPLDMSLSTHSSVQAGGWRKLPPLNLSPTIEYNDG 2520  
 Db 2460 SFESLSPSSRPASPTRQAOQTPVLSPLDMSLSTHSSVQAGGWRKLPPLNLSPTIEYNDG 2519  
 QY 2521 RPAKRHDIARSHSSESPRLPINRSGTWKRHSKHSLSPLRVSTWRTGSSSSILSASES 2580  
 Db 2520 RPAKRHDIARSHSSESPRLPINRSGTWKRHSKHSLSPLRVSTWRTGSSSSILSASES 2579  
 QY 2581 SEKAKSEDEKHVNSISGTQKSENOVSAKGTWRKI KENEFSPNTNSTQTVSSGATNGAES 2640  
 Db 2580 SEKAKSEDEKHVNSISGTQKSENOVSAKGTWRKI KENEFSPNTNSTQTVSSGATNGAES 2639  
 QY 2641 KTLIYQMAPAVSKTEDVWVRIEDCPINNPRSGRSPGTGNTPPVIDSVSEKAMPNIKDSKDN 2700  
 Db 2640 KTLIYQMAPAVSKTEDVWVRIEDCPINNPRSGRSPGTGNTPPVIDSVSEKAMPNIKDSKDN 2699  
 QY 2701 QAKQNVGNGSVPMETVGLNRLTSFIQVDAPOKGTETIKPGONNPVPVSETNESPIVERT 2760  
 Db 2700 QAKQNVGNGSVPMETVGLNRLTSFIQVDAPOKGTETIKPGONNPVPVSETNESPIVERT 2759  
 QY 2761 PFSSSSSSSKHSSPSGTVAAARVTPFNYPNPSPRKSSADSTSARPSQIPTPVNNNTKKRDKT 2820  
 Db 2760 PFSSSSSSSKHSSPSGTVAAARVTPFNYPNPSPRKSSADSTSARPSQIPTPVNNNTKKRDKT 2819  
 QY 2821 DSTSSSGTQSPKRHSGSYLVTSV 2843  
 Db 2820 DSTSSSGTQSPKRHSGSYLVTSV 2842

Search completed: August 25, 2004, 17:14:55  
 Job time : 198.5 secs

**This Page Blank (uspio)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2004, 17:23:12 ; Search time 190.5 Seconds  
(without alignments)  
4895.247 Million call updates/sec

Title: US-09-442-489F-2

Perfect score: 14575

Sequence: 1 MAASVDQLKQVEALKMEN.....ESSGTQPKRHSGSYLTVTSV 2843

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1297172 seqs, 314612898 residues

Total number of hits satisfying chosen parameters: 1297172

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*

2: /cgn2\_6/ptodata/2/pubpaa/FCT\_NEW\_PUB.pep.\*

3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*

4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*

5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*

6: /cgn2\_6/ptodata/2/pubpaa/FCTUS\_PUBCOMB.pep.\*

7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*

8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*

9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*

10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*

11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*

12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*

13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*

14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*

15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*

16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*

17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*

18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	14533	99.7	2843	9	US-09-987-482-1
2	14533	99.7	2843	12	US-10-392-113-21
3	14526.5	99.7	2844	12	US-10-267-502-370
4	14524	99.7	2843	16	US-10-408-765A-1970
5	14506	99.5	2843	8	US-08-681-219-32
6	14506	99.5	2843	11	US-09-230-111C-30
7	14506	99.5	2843	14	US-10-092-138-30
8	13148	90.2	2845	12	US-10-267-502-372
9	4719	32.4	912	9	US-09-987-482-2
10	3889	26.7	767	9	US-09-987-482-3
11	3508.5	24.1	2274	12	US-10-267-502-373
12	3414.5	23.4	2303	12	US-10-267-502-371
13	1763	12.1	2417	12	US-10-267-502-368
14	1640	11.3	325	12	US-09-915-307-5
15	1474	10.1	1633	14	US-10-029-386-33090

Query Match 99.7%; Score 14533; DB 9; Length 2843;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2836; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy	1	MAASVDQLKQVEALKMENLNRLQELNHNLTKEASNNKVKLQQLQSSIDEAM	60
Db	1	MAASVDQLKQVEALKMENLNRLQELNHNLTKEASNNKVKLQQLQSSIDEAM	60
Qy	61	ASSGQDILLERLKLNDSSNFPQVKLRKXMSLRYSGRSGSVSRSGECSVPVMSGFPFR	120
Db	61	ASSGQDILLERLKLNDSSNFPQVKLRKXMSLRYSGRSGSVSRSGECSVPVMSGFPFR	120
Qy	121	RGFVNGSRESGYLEELKEKRSLLADLDKEEKEKDWYQAQLQNLTKRIDSLPTENFSL	180
Db	121	RGFVNGSRESGYLEELKEKRSLLADLDKEEKEKDWYQAQLQNLTKRIDSLPTENFSL	180
Qy	191	QTDLTRCLLEVEARQIRVAMEEQIGTQDMKRAQRRIATQQLKDLIRLROLQSOAT	240

## ALIGNMENTS

### RESULT 1

US-09-987-482-1  
; Sequence 1, Application US/09987482  
; Publication No. US20020184656A1  
; GENERAL INFORMATION:  
; APPLICANT: BHANDARI, POONAM  
; APPLICANT: SHASHIDHARA, L.S.  
; TITLE OF INVENTION: IN VIVO ASSAY SYSTEM FOR SCREENING AND VALIDATION OF  
; TITLE OF INVENTION: DRUGS AND OTHER SUBSTANCES  
; FILE REFERENCE: 056859-0134  
; CURRENT APPLICATION NUMBER: US/09/987,482  
; CURRENT FILING DATE: 2002-03-21  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 2843  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-987-482-1

Sequence 369, Appl  
Sequence 12713, A  
Sequence 2231, Ap  
Sequence 43924, A  
Sequence 4, Appli  
Sequence 49, Appli  
Sequence 104793,  
Sequence 142, App  
Sequence 70580, A  
Sequence 23, Appli  
Sequence 5, Appli  
Sequence 5784, Ap  
Sequence 143, App  
Sequence 423, App  
Sequence 2287, Ap  
Sequence 73678, A  
Sequence 144, App  
Sequence 83, Appli  
Sequence 71033, A  
Sequence 34248, A  
Sequence 11, Appli  
Sequence 8, Appli  
Sequence 5, Appli  
Sequence 71235, A  
Sequence 62, Appli  
Sequence 81, Appli  
Sequence 359, App  
Sequence 12, Appli  
Sequence 8, Appli

16 1221.5 8.4 1067 12 US-10-267-502-369  
17 556.5 3.8 2344 9 US-09-815-242-12713  
18 536 3.7 2263 16 US-10-408-765A-2231  
19 518 3.6 2271 12 US-10-282-122A-43924  
20 517.5 3.6 2283 14 US-10-172-502-4  
21 507.5 3.5 178 13 US-10-071-751-49  
22 503 3.5 4322 16 US-10-437-963-104793  
23 474.5 3.3 5560 12 US-10-263-929-142  
24 461.5 3.2 6641 12 US-10-282-122A-70580  
25 460.5 3.2 10203 16 US-10-661-809-23  
26 455.5 3.1 22152 16 US-10-715-066-5  
27 452 3.1 3507 15 US-10-369-493-5784  
28 445 3.1 3664 12 US-10-263-929-143  
29 445 3.1 3664 14 US-10-177-293-43  
30 445 3.1 3664 16 US-10-408-765A-2287  
31 442 3.0 4776 12 US-10-282-122A-73678  
32 436.5 3.0 3551 12 US-10-263-929-144  
33 433.5 3.0 2701 14 US-10-171-311-83  
34 431.5 3.0 1831 12 US-10-282-122A-71033  
35 429 2.9 2665 9 US-09-864-761-34248  
36 429 2.9 5877 14 US-10-142-515-11  
37 416 2.9 4723 15 US-10-359-012-8  
38 414 2.8 5935 14 US-10-243-243A-8  
39 409.5 2.8 2803 12 US-10-415-187-5  
40 408 2.8 3692 12 US-10-282-122A-71235  
41 403.5 2.8 3051 15 US-10-144-194A-62  
42 398 2.7 2137 12 US-10-042-865-81  
43 398 2.7 2432 12 US-10-112-944-359  
44 398 2.7 2545 12 US-10-042-865-12  
45 395 2.7 2429 12 US-10-362-892-8

Db 181 QTDNTRQLEYEARQIRVAMEEQJGTQCDMEKRAQRRIAQIIEKDILIRIOLLOSOAT 240  
Qy 241 EAERSSQNKHETGSHDAERQNEGQVGEINMATSNGQGGTTRMDHETASVLSSSTHSA 300  
Db 241 EAERSSQNKHETGSHDAERQNEGQVGEINMATSNGQGGTTRMDHETASVLSSSTHSA 300  
Qy 301 PRRLTSHLGTKEVMYSLLSNGLGTHDXDMSRTLAMSSQDSCISNRQSGCPLLIQLL 360  
Db 301 PRRLTSHLGTKEVMYSLLSNGLGTHDXDMSRTLAMSSQDSCISNRQSGCPLLIQLL 360  
Qy 361 HGNDKDSVLLGNSRGSKEARASAAALHNIHSQDDKGRREIRVLHLEQIRAYCETC 420  
Db 361 HGNDKDSVLLGNSRGSKEARASAAALHNIHSQDDKGRREIRVLHLEQIRAYCETC 420  
Qy 421 WEWQEAHEPGMDQDNMPAPVEHQICPAVCVLMKLSFDEEHRHAMNELGGLQAIABELLQ 480  
Db 421 WEWQEAHEPGMDQDNMPAPVEHQICPAVCVLMKLSFDEEHRHAMNELGGLQAIABELLQ 480  
Qy 481 VDCEMYGLTNDHYSITLRRYAGMALTNLTFGDVANKATLCMSKGCYRALVAQILKSEEDL 540  
Db 481 VDCEMYGLTNDHYSITLRRYAGMALTNLTFGDVANKATLCMSKGCYRALVAQILKSEEDL 540  
Qy 541 QQVITASVLRNLMSRADVNSKKTREVGSVKALMECALEVKESTLKSALWNLGAHCT 600  
Db 541 QQVITASVLRNLMSRADVNSKKTREVGSVKALMECALEVKESTLKSALWNLGAHCT 600  
Qy 601 ENKADICAVDGAFLVGTLYTYSQNTLAIIESGGGILRWSSLIATNEDHROIIRENN 660  
Db 601 ENKADICAVDGAFLVGTLYTYSQNTLAIIESGGGILRWSSLIATNEDHROIIRENN 660  
Qy 661 CLQTLLOHLKSHSITIVSNACGLTWNLSARNPKDQALWDMGAVMLKNLIHSHKHWMIA 720  
Db 661 CLQTLLOHLKSHSITIVSNACGLTWNLSARNPKDQALWDMGAVMLKNLIHSHKHWMIA 720  
Qy 721 GSAAALRNLMANRPKADANIMSPGSLPSLHVRRKQKALEABEADQHLSETFDNDNL 780  
Db 721 GSAAALRNLMANRPKADANIMSPGSLPSLHVRRKQKALEABEADQHLSETFDNDNL 780  
Qy 781 PKASHRSKQHKQSLYGDYVDTNRHDDNDSFNFTGNMTVLPYNTTTLVLPSSSSRGS 840  
Db 781 PKASHRSKQHKQSLYGDYVDTNRHDDNDSFNFTGNMTVLPYNTTTLVLPSSSSRGS 840  
Qy 841 LSSRSSEKRSLEMERGIGLGNHPATENPGTSKRGLOISTTAAGIAKVMESVAIHTS 900  
Db 841 LSSRSSEKRSLEMERGIGLGNHPATENPGTSKRGLOISTTAAGIAKVMESVAIHTS 900  
Qy 901 QEDRSSGTTTELHCVTDERALNRSSAAHSHNTYNTFKSENNRCTSPYAKLEYKRSS 960  
Db 901 QEDRSSGTTTELHCVTDERALNRSSAAHSHNTYNTFKSENNRCTSPYAKLEYKRSS 960  
Qy 961 NDSLNSVSSNDGYKRGOMKPSIESYEDDESFCVGYQYPADLAHKIHSANHMDNDGE 1020  
Db 961 NDSLNSVSSNDGYKRGOMKPSIESYEDDESFCVGYQYPADLAHKIHSANHMDNDGE 1020  
Qy 1021 LDTPTNLSKYSDBLNSGRQSPONERWAPKHIIIEDEIKQSEQRQSNQSTTYPVYTE 1080  
Db 1021 LDTPTNLSKYSDBLNSGRQSPONERWAPKHIIIEDEIKQSEQRQSNQSTTYPVYTE 1080  
Qy 1081 STDDKHLKFPQHQBCVSPYRSRGANGSETNRVSGNHGINVWSOLCOEDDYEDDKP 1140  
Db 1081 STDDKHLKFPQHQBCVSPYRSRGANGSETNRVSGNHGINVWSOLCOEDDYEDDKP 1140  
Qy 1141 TNSYERYSEBQHEEBERPTNYSIKYNEEKRVDPQIDYSLKYATDIPSSQKQSFSSKS 1200  
Db 1141 TNSYERYSEBQHEEBERPTNYSIKYNEEKRVDPQIDYSLKYATDIPSSQKQSFSSKS 1200  
Qy 1201 SSGOSSKTEHMSSESNTSTPSSNAKQNLHPSSAQSRSGQPKAATCKVSSINQETIQ 1260  
Db 1201 SSGOSSKTEHMSSESNTSTPSSNAKQNLHPSSAQSRSGQPKAATCKVSSINQETIQ 1260  
Qy 1261 TYCVEDTPICFSRCSLSLSAABDEICGNCQTTQEADPSANTLQIAIEKIGKIGTRSAEDPV 1320  
Db 1261 TYCVEDTPICFSRCSLSLSAABDEICGNCQTTQEADPSANTLQIAIEKIGKIGTRSAEDPV 1320

Qy 1321 SEVPAVSQHPRTKSRLOGSSLSSESARHKAVERPPSGAKSPSKGAOTPKPSPPHYQET 1380  
Db 1321 SEVPAVSQHPRTKSRLOGSSLSSESARHKAVERPPSGAKSPSKGAOTPKPSPPHYQET 1380  
Qy 1381 PLMPFRCTSVSSLSDFESRSIASSVQSEPCSGMVSGIISPSDLDPSPGQOTWPPRSKTPP 1440  
Db 1381 PLMPFRCTSVSSLSDFESRSIASSVQSEPCSGMVSGIISPSDLDPSPGQOTWPPRSKTPP 1440  
Qy 1441 PPQTAQTKREVKNKAPTAEKRESGPKQAANAARVQVLPADATLLHFATESTPDGF 1500  
Db 1441 PPQTAQTKREVKNKAPTAEKRESGPKQAANAARVQVLPADATLLHFATESTPDGF 1500  
Qy 1501 SCSSLSALSLEDPFIQKDVLRIMPVQENDNGNETESEQPKESNENQEKAEKTIIDSE 1560  
Db 1501 SCSSLSALSLEDPFIQKDVLRIMPVQENDNGNETESEQPKESNENQEKAEKTIIDSE 1560  
Qy 1561 KDLDDSDDDDIIELEECIIISAMPTKSRKKGKPAQTASKLPPPVARKPSOLPVYKLLPS 1620  
Db 1561 KDLDDSDDDDIIELEECIIISAMPTKSRKKGKPAQTASKLPPPVARKPSOLPVYKLLPS 1620  
Qy 1621 QNRLQPKXVSFTPGDDMPRVYCVEGTPIINFSTATSLSDLTIESPPNELAAGEGVGGAQ 1680  
Db 1621 QNRLQPKXVSFTPGDDMPRVYCVEGTPIINFSTATSLSDLTIESPPNELAAGEGVGGAQ 1680  
Qy 1681 SGEFEKRTDITPTEGRSTDEAOGKTSVVITPELDONKAEEDGIIAECINSAMPKGSHP 1740  
Db 1681 SGEFEKRTDITPTEGRSTDEAOGKTSVVITPELDONKAEEDGIIAECINSAMPKGSHP 1740  
Qy 1741 FRVKIMDQVQOASASSAPNKNQLDGKKKXPTSPVKPIQONTYRTRVRKNADSKNIN 1800  
Db 1741 FRVKIMDQVQOASASSAPNKNQLDGKKKXPTSPVKPIQONTYRTRVRKNADSKNIN 1800  
Qy 1801 ABRVFSNDKSKQNLKNNKDFNDKLPNNEDVRGSAFDSPHYTPIEGTTCFSEND 1860  
Db 1801 ABRVFSNDKSKQNLKNNKDFNDKLPNNEDVRGSAFDSPHYTPIEGTTCFSEND 1860  
Qy 1861 SLSSLDFODDDDVLSREKAEELRKAKENKESAKVTSTELTSTNOOSANKTQAIKQPINR 1920  
Db 1861 SLSSLDFODDDDVLSREKAEELRKAKENKESAKVTSTELTSTNOOSANKTQAIKQPINR 1920  
Qy 1921 GQPKPILQKQSTFPQSSKQIDPRGAATDEKLONPAIENTPVCFSHNSLSLSIDSEDLQ 1980  
Db 1921 GQPKPILQKQSTFPQSSKQIDPRGAATDEKLONPAIENTPVCFSHNSLSLSIDSEDLQ 1980  
Qy 1981 NKENEPIKETPPSQCEPSPKQAGVAPKSFHVEDTPVCFSRNSSLSLSIDSEDLQ 2040  
Db 1981 NKENEPIKETPPSQCEPSPKQAGVAPKSFHVEDTPVCFSRNSSLSLSIDSEDLQ 2040  
Qy 2041 ECISAMPKKKXKRLKGDNEKHSFRNMGIGLEDLTLDKQIORPSEHGLSPDSENF 2100  
Db 2041 ECISAMPKKKXKRLKGDNEKHSFRNMGIGLEDLTLDKQIORPSEHGLSPDSENF 2100  
Qy 2101 WKATOEGANSIVSSLHQAAAAACLSROASDSDSLKSGISLGSFHLTPDOEEKPFT 2160  
Db 2101 WKATOEGANSIVSSLHQAAAAACLSROASDSDSLKSGISLGSFHLTPDOEEKPFT 2160  
Qy 2161 SNKGPRIILKPEKSTLETTKTIESESKGKGGKVVYKSLITGKVRNSSEISQMKQPLQAN 2220  
Db 2161 SNKGPRIILKPEKSTLETTKTIESESKGKGGKVVYKSLITGKVRNSSEISQMKQPLQAN 2220  
Qy 2221 MPSISRGHTMTIHIFGVNRSSSSTSPVKKGPPLKTPASKSPSECOQTATTPRGAKPSVK 2280  
Db 2221 MPSISRGHTMTIHIFGVNRSSSSTSPVKKGPPLKTPASKSPSECOQTATTPRGAKPSVK 2280  
Qy 2281 ELSPVARTSIOIGSSKAPSRSGSRDSTPSRPAQPLSRPIQSPGRNSISPGRNGISPPN 2340  
Db 2281 ELSPVARTSIOIGSSKAPSRSGSRDSTPSRPAQPLSRPIQSPGRNSISPGRNGISPPN 2340  
Qy 2341 KLSQLPRTSSTASTKSSGSGKSYTSPGRMSQNLTKQTGLSKNASSIPRSESASK 2400  
Db 2341 KLSQLPRTSSTASTKSSGSGKSYTSPGRMSQNLTKQTGLSKNASSIPRSESASK 2400



```
QY 2401 LNMNNGANKKVELSRMSSKSGESDRSRRPVLVRQSTFIKEAPSPTLRRKLEBSA 2460
DB 2401 LNMNNGANKKVELSRMSSKSGESDRSRRPVLVRQSTFIKEAPSPTLRRKLEBSA 2460
QY 2461 SPESLSPSRPASPTRSQATPVLSPSLPDMGLSTHSSVQAGWRKLPNLSPTTIEYNDG 2520
DB 2461 SPESLSPSRPASPTRSQATPVLSPSLPDMGLSTHSSVQAGWRKLPNLSPTTIEYNDG 2520
QY 2521 RPAKHDIARSHSESPSRPIPNRSCTWKKEHSHKSSIPRVSTWRTGSSSILSASSES 2580
DB 2521 RPAKHDIARSHSESPSRPIPNRSCTWKKEHSHKSSIPRVSTWRTGSSSILSASSES 2580
QY 2581 SEKAKSEDEKHVNSISGTRQSKENOVSAKGTWRKIKENEFSPTNSTQTSVSSGATNGAES 2640
DB 2581 SEKAKSEDEKHVNSISGTRQSKENOVSAKGTWRKIKENEFSPTNSTQTSVSSGATNGAES 2640
QY 2641 KTLIYQMAPAVSKTEDVWRIEDCPINNPRSGRSPGTNTPPVIDSVSEKANPINIKSDKN 2700
DB 2641 KTLIYQMAPAVSKTEDVWRIEDCPINNPRSGRSPGTNTPPVIDSVSEKANPINIKSDKN 2700
QY 2701 QAKQVNGSGVPMRTVGLNRLTSTFIQVDAPQKGTETKPGQNNPVVSETNESPIVERT 2760
DB 2701 QAKQVNGSGVPMRTVGLNRLTSTFIQVDAPQKGTETKPGQNNPVVSETNESPIVERT 2760
QY 2761 PFSSSSSSKHSSPFGTVAARVTPFNPNPSRKSSADTSARPISOIPTPVNNNTKKRDSKT 2820
DB 2761 PFSSSSSSKHSSPFGTVAARVTPFNPNPSRKSSADTSARPISOIPTPVNNNTKKRDSKT 2820
QY 2821 DSTESGTOCPKXHGSGYLVTSV 2843
DB 2821 DSTESGTOCPKXHGSGYLVTSV 2843

RESULT 2
US-10-392-113-21
; Sequence 21, Application US/10392113
; Publication No. US2003024993A1
; GENERAL INFORMATION:
; APPLICANT: Land, Laurent
; TITLE OF INVENTION: COMPOSITIONS THAT INHIBIT PROLIFERATION
; FILE REFERENCE: 21108.0005U3
; CURRENT APPLICATION NUMBER: US/10/392,113
; PRIORITY FILING DATE: 2003-03-17
; PRIOR APPLICATION NUMBER: 60/365,078
; PRIORITY FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: PCT/US01/32127
; PRIORITY FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 60/239,705
; PRIORITY FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 2843
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:/Note =
; OTHER INFORMATION: Synthetic Construct
US-10-392-113-21

Query Match 99.7%; Score 14533; DB 12; Length 2843;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2836; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 MAAAYDQLLKQVEALKMENSLNRQLEDNLSNHLTKLETEASNKMEVLKQLQGSIEDEAM 60
DB 1 MAAAYDQLLKQVEALKMENSLNRQLEDNLSNHLTKLETEASNKMEVLKQLQGSIEDEAM 60
QY 61 ASSGQDILLERLKEINLDSSNFFGVKLRKMSLRSGVSGREGSVSRGSCSPVPMGSPFR 120
DB 61 ASSGQDILLERLKEINLDSSNFFGVKLRKMSLRSGVSGREGSVSRGSCSPVPMGSPFR 120
```

```
QY 121 RGFVNGSRESTGYLLELEKERSLILADLDKEBEKDWYAOQLNLTKRIIDSLPLTENFSL 180
DB 121 RGFVNGSRESTGYLLELEKERSLILADLDKEBEKDWYAOQLNLTKRIIDSLPLTENFSL 180
QY 181 QTDLTRQLRYEARQIRVAMEEQIGTCCDMEKBAQRIARIQOITEKIDILRILQLOSOAT 240
DB 181 QTDLTRQLRYEARQIRVAMEEQIGTCCDMEKBAQRIARIQOITEKIDILRILQLOSOAT 240
QY 241 EARSSEKQKHETGSHDAERONEGQGVGEINWATSGNGCGSTTRMDHETASVLSSSSTHSA 300
DB 241 EARSSEKQKHETGSHDAERONEGQGVGEINWATSGNGCGSTTRMDHETASVLSSSSTHSA 300
QY 301 PRRLTSHLGTKEVMVYLLSMLGTHDKDDMSRTLLAMSSSQDSCISMROGCGCLPLLIQLL 360
DB 301 PRRLTSHLGTKEVMVYLLSMLGTHDKDDMSRTLLAMSSSQDSCISMROGCGCLPLLIQLL 360
QY 361 HGNDKDSVLLGNSRSGSEARASAAALHNTIHSQDDKRGRRREIRVUHLLEQIRAYCETC 420
DB 361 HGNDKDSVLLGNSRSGSEARASAAALHNTIHSQDDKRGRRREIRVUHLLEQIRAYCETC 420
QY 421 WEMQEAHEPGMDQDNMPAPVEHOICPAVCVLMKLSFDEEHRHAMNELGGLQIAIELLO 480
DB 421 WEMQEAHEPGMDQDNMPAPVEHOICPAVCVLMKLSFDEEHRHAMNELGGLQIAIELLO 480
QY 481 VDCMYGLTNDHYSITLRRYAGMALTNLTFGDVANKATLCMKGCMRALVAQLKSESEDL 540
DB 481 VDCMYGLTNDHYSITLRRYAGMALTNLTFGDVANKATLCMKGCMRALVAQLKSESEDL 540
QY 541 QOVTASVLRNLSWRADVNSKKTIREVGSVKALMECALEVKESTLKVLSALMNLASHCT 600
DB 541 QOVTASVLRNLSWRADVNSKKTIREVGSVKALMECALEVKESTLKVLSALMNLASHCT 600
QY 601 ENKADICAVDGAFLVGLTLYRSQTNLTALIESGGGILRNVSLSIATNEDHROILRENN 660
DB 601 ENKADICAVDGAFLVGLTLYRSQTNLTALIESGGGILRNVSLSIATNEDHROILRENN 660
QY 661 CLQTLLOHLKXSLTIVSNACGTLWNLARNPKDQEALEWDMGAVMLKXLIHSHKMIAM 720
DB 661 CLQTLLOHLKXSLTIVSNACGTLWNLARNPKDQEALEWDMGAVMLKXLIHSHKMIAM 720
QY 721 GSAALRNLMANRPAYKDNIMSPGSSLPSLHVRKQKALEAEADAOHSETFNDINLS 780
DB 721 GSAALRNLMANRPAYKDNIMSPGSSLPSLHVRKQKALEAEADAOHSETFNDINLS 780
QY 781 PKASHRSKQRHKQSLYGDYVFDNRRHDDNRSDNFTGNMTVLSPLYNTTVLPSSSSSRGS 840
DB 781 PKASHRSKQRHKQSLYGDYVFDNRRHDDNRSDNFTGNMTVLSPLYNTTVLPSSSSSRGS 840
QY 841 LDSSRSKDRSLERERGIIGLVNHPATENPCTSSKRGLOJSTTAAQAKWEEVSAIHTS 900
DB 841 LDSSRSKDRSLERERGIIGLVNHPATENPCTSSKRGLOJSTTAAQAKWEEVSAIHTS 900
QY 901 QEDRSSGSTTELHCVTDERNALRRSSAAHSTNFTKSENSNRCTCMPYAKLEYKRSS 960
DB 901 QEDRSSGSTTELHCVTDERNALRRSSAAHSTNFTKSENSNRCTCMPYAKLEYKRSS 960
QY 961 NDSLNSVSSNDGYKRGOMKPSIESYEDDESKECSYGOVPADLAHKIHSANHMDNDGE 1020
DB 961 NDSLNSVSSNDGYKRGOMKPSIESYEDDESKECSYGOVPADLAHKIHSANHMDNDGE 1020
QY 1021 LDTPTINYSKYSDEQLNSGROSPQNERWARPKHIIIEDEIKQSEQRORNRQSTTYPVYTE 1080
DB 1021 LDTPTINYSKYSDEQLNSGROSPQNERWARPKHIIIEDEIKQSEQRORNRQSTTYPVYTE 1080
QY 1081 STDDKHLKQFHFQGCQECVSPYRGANGSETNRVGNHGINQVNSOSLCOEDDYEDDKP 1140
DB 1081 STDDKHLKQFHFQGCQECVSPYRGANGSETNRVGNHGINQVNSOSLCOEDDYEDDKP 1140
QY 1141 TNSYERISEEQQHEEERPTNYSIKYNEEKHVDQPIDYSLKYATDIPSSQKQSFSPKS 1200
DB 1141 TNSYERISEEQQHEEERPTNYSIKYNEEKHVDQPIDYSLKYATDIPSSQKQSFSPKS 1200
```

QY 1201 SSGSSKTEHSSSSSENTSTPSSNAKQNLHPSSAQSRSQGPOKAATCKVSSINQETIQ 1260  
 Db 1201 SSGSSKTEHSSSSSENTSTPSSNAKQNLHPSSAQSRSQGPOKAATCKVSSINQETIQ 1260  
 QY 1261 TYCVEDTPICFSRGSSLSLSAASDEIGCNOTTOEADSANTLQIAEIKGKIGTRSAEDPV 1320  
 Db 1261 TYCVEDTPICFSRGSSLSLSAASDEIGCNOTTOEADSANTLQIAEIKGKIGTRSAEDPV 1320  
 QY 1321 SEVPAVSOPHRTKSRRLQGSLSSESARHKAVERPPSGAKSPSKCAOTPKSPPEHYVOET 1380  
 Db 1321 SEVPAVSOPHRTKSRRLQGSLSSESARHKAVERPPSGAKSPSKCAOTPKSPPEHYVOET 1380  
 QY 1381 PLMFSRCTSVSLDSFESRSIASVQSEPCSGMVSGIISPSDLDPSPGQTMPFSRSKTPP 1440  
 Db 1381 PLMFSRCTSVSLDSFESRSIASVQSEPCSGMVSGIISPSDLDPSPGQTMPFSRSKTPP 1440  
 QY 1441 PPPQTAOTKREVPKNKAPTAEKRESGPKQAANAAVORVVLDPADTLLHFAIESTPDGF 1500  
 Db 1441 PPPQTAOTKREVPKNKAPTAEKRESGPKQAANAAVORVVLDPADTLLHFAIESTPDGF 1500  
 QY 1501 SCSSLSALSDEPFIQXDELVRIMPVPOENDNGNETESEOPKESNENQKEAEKTIIDSE 1560  
 Db 1501 SCSSLSALSDEPFIQXDELVRIMPVPOENDNGNETESEOPKESNENQKEAEKTIIDSE 1560  
 QY 1561 KDLDDSDDDIEILKECIIISAMPTKSRGKKAQATASKLPPVARKPSQLPVYKLLPS 1620  
 Db 1561 KDLDDSDDDIEILKECIIISAMPTKSRGKKAQATASKLPPVARKPSQLPVYKLLPS 1620  
 QY 1621 QNRQPQKHVSFTPGDDMPRVYCVGEGTPIINFSTATSLSDLTIESPPNELAAGEVRCGAQ 1680  
 Db 1621 QNRQPQKHVSFTPGDDMPRVYCVGEGTPIINFSTATSLSDLTIESPPNELAAGEVRCGAQ 1680  
 QY 1681 SGEFEKRTIPIEGRSTDEAGGKTSSVTPELDNDKAEGDIIAECINSAMPKGSHPK 1740  
 Db 1681 SGEFEKRTIPIEGRSTDEAGGKTSSVTPELDNDKAEGDIIAECINSAMPKGSHPK 1740  
 QY 1741 FRVKIMDQVOQAASASAPNNKQLDGKKKPTSPVKPIPONTYRTRVRKNADSKNNLN 1800  
 Db 1741 FRVKIMDQVOQAASASAPNNKQLDGKKKPTSPVKPIPONTYRTRVRKNADSKNNLN 1800  
 QY 1801 AERFSDNKSKQNLKNNKDFNDKLPNNEDRVGSGFAFDSPHHYTPIEGTPVCFSRND 1860  
 Db 1801 AERFSDNKSKQNLKNNKDFNDKLPNNEDRVGSGFAFDSPHHYTPIEGTPVCFSRND 1860  
 QY 1861 SLSSLDPDDDDVLSREKAEILKAKENKESAKVTSHTELTSNCOASANKQTAQKQPINR 1920  
 Db 1861 SLSSLDPDDDDVLSREKAEILKAKENKESAKVTSHTELTSNCOASANKQTAQKQPINR 1920  
 QY 1921 GQPKPILQKSTFPQSSXDIIDRGAATDEKLQNFPAIENTPVCFSHNSLSLSLSDIDQENN 1980  
 Db 1921 GQPKPILQKSTFPQSSXDIIDRGAATDEKLQNFPAIENTPVCFSHNSLSLSLSDIDQENN 1980  
 QY 1981 NKENPPIKETPPSQSGPFPQASGAPKSFHVEDTPVCFSRNSSLSSLSISDEDDLQ 2040  
 Db 1981 NKENPPIKETPPSQSGPFPQASGAPKSFHVEDTPVCFSRNSSLSSLSISDEDDLQ 2040  
 QY 2041 ECISAMPKPKKPSRLKGDNEKHSFRNMGGLIGBDLTLDLKDIOQRPDSEHGLSPDSNF 2100  
 Db 2041 ECISAMPKPKKPSRLKGDNEKHSFRNMGGLIGBDLTLDLKDIOQRPDSEHGLSPDSNF 2100  
 QY 2101 WKATQEGANGIVSSLHQAAAAACLSRQASSDSISLSKGISLGSPFHLTPDQEEKPFT 2160  
 Db 2101 WKATQEGANGIVSSLHQAAAAACLSRQASSDSISLSKGISLGSPFHLTPDQEEKPFT 2160  
 QY 2161 SNKGPRILKPEKSTLETFTKTESSEKGIKGGKVVYKSLITGKVRNSBEISGQMKQPLQAN 2220  
 Db 2161 SNKGPRILKPEKSTLETFTKTESSEKGIKGGKVVYKSLITGKVRNSBEISGQMKQPLQAN 2220  
 QY 2221 MPFSIRGRMTIHIPVRNNSSTSPVSKGPPKLTTPASKPSPEGQATTTPRGAKPSVK 2280  
 Db 2221 MPFSIRGRMTIHIPVRNNSSTSPVSKGPPKLTTPASKPSPEGQATTTPRGAKPSVK 2280  
 QY 2281 ELSPVARQTSQIGSSKAPSRSGSRDSTPSRPAQQLSRPIQSPGRNSISPGRNGISPPN 2340

Db 2281 ELSPVARQTSQIGSSKAPSRSGSRDSTPSRPAQQLSRPIQSPGRNSISPGRNGISPPN 2340  
 QY 2341 KLSQIPRSTSPSTASTKSSGSKMSYTPGQMSQNLTKOTGLSKNASSIIPRESASKG 2400  
 Db 2341 KLSQIPRSTSPSTASTKSSGSKMSYTPGQMSQNLTKOTGLSKNASSIIPRESASKG 2400  
 QY 2401 LNQMNGANGANKVLSRMSSTKSSGESDESERFVLVRQSTFIKEAPSPTLRKLLESA 2460  
 Db 2401 LNQMNGANGANKVLSRMSSTKSSGESDESERFVLVRQSTFIKEAPSPTLRKLLESA 2460  
 QY 2461 SFESLSPSRPASPTRSQAOTPVLSPLPDMSLSTHSSVQAGGWRKLPNLSPTIEYNDG 2520  
 Db 2461 SFESLSPSRPASPTRSQAOTPVLSPLPDMSLSTHSSVQAGGWRKLPNLSPTIEYNDG 2520  
 QY 2521 RPAKRDHARSSESFSLPNNRSGTWKREHSKHSSSLPRVSTWRTGSSSSIISSASSES 2580  
 Db 2521 RPAKRDHARSSESFSLPNNRSGTWKREHSKHSSSLPRVSTWRTGSSSSIISSASSES 2580  
 QY 2581 SEKAKSEDEKHVNSISGTSKQKNOVSAGTWRKI KENEFPTNSTSTQTVSSGATNGAES 2640  
 Db 2581 SEKAKSEDEKHVNSISGTSKQKNOVSAGTWRKI KENEFPTNSTSTQTVSSGATNGAES 2640  
 QY 2641 KTLIYQMAPAVSKTDDVVRIEDCPINNPRGSRPTGNTPPVIDSVSEKANPNIKDSKN 2700  
 Db 2641 KTLIYQMAPAVSKTDDVVRIEDCPINNPRGSRPTGNTPPVIDSVSEKANPNIKDSKN 2700  
 QY 2701 QAKQNVGSGVPMRTVGLNRLTSTFIQVADAPQKTEIKPQONNPVPVSETNESPIVERT 2760  
 Db 2701 QAKQNVGSGVPMRTVGLNRLTSTFIQVADAPQKTEIKPQONNPVPVSETNESPIVERT 2760  
 QY 2761 PSSSSSSKHSGSPCTVAARVTPNNPNSPRKSSADTSASPSQIPTVNNNTKRDSDKT 2820  
 Db 2761 PSSSSSSKHSGSPCTVAARVTPNNPNSPRKSSADTSASPSQIPTVNNNTKRDSDKT 2820  
 QY 2821 DSTESSGTQSPKRHSGSYLVTSV 2843  
 Db 2821 DSTESSGTQSPKRHSGSYLVTSV 2843  
 RESULT 3  
 US-10-267-502-370  
 ; Sequence 370, Application US/10267502  
 ; Publication No. US20040071700A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kim, Jaeseob  
 ; TITLE OF INVENTION: Obesity Linked Genes  
 ; FILE REFERENCE: LSD-07416  
 ; CURRENT APPLICATION NUMBER: US/10/267,502  
 ; CURRENT FILING DATE: 2003-01-27  
 ; NUMBER OF SEQ ID NOS: 439  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 370  
 ; LENGTH: 2844  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-267-502-370  
 Query Match 99.7%; Score 14526.5; DB 12; Length 2844;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 2837; Conservative 1; Mismatches 5; Indels 1; Gaps 1;  
 QY 1 MAASVDQLLKQVEALKKMNENLRLQEDNLSNHLTKLETSANMKVVKLQKQSGIEDAM 60  
 Db 1 MAASVDQLLKQVEALKKMNENLRLQEDNLSNHLTKLETSANMKVVKLQKQSGIEDAM 60  
 QY 61 ASSGQIDLLERLKEINLDSNPPGVKLSKMSLSRSGVSGREGSVSRSGCECSPVPMGSPFR 120  
 Db 61 ASSGQIDLLERLKEINLDSNPPGVKLSKMSLSRSGVSGREGSVSRSGCECSPVPMGSPFR 120  
 QY 121 RGTVNGSRESTGYLEBELEKERSLLIADLDKBEKMDYTAQQLNLTKRIDSILPUTENFSL 180

Db 121 RGFVNGSRESTGYLEELEKERSLILADLDKEEKEKDWYAOQLNLTXRIDSLPLTENFSL 180  
Qy 181 QTDLTRRQLEYEARQIRVAMEEQLTQODMEKRAORRIARIQOITEKDILIRIQLLOQAT 240  
Db 181 QTDLTRRQLEYEARQIRVAMEEQLTQODMEKRAORRIARIQOITEKDILIRIQLLOQAT 240  
Qy 241 EAEKSSQNHKTGSHDAERQEGGVGEINWATSGNGGSTRVDHETASVLSSSSTHSA 300  
Db 241 EAEKSSQNHKTGSHDAERQEGGVGEINWATSGNGGSTRVDHETASVLSSSSTHSA 300  
Qy 301 PRRLTSHLGTKEVWYVYSLLSMLGTHDKDDMSRTLLAMSSQDSCISMRQSGCLPLLIQLL 360  
Db 301 PRRLTSHLGTKEVWYVYSLLSMLGTHDKDDMSRTLLAMSSQDSCISMRQSGCLPLLIQLL 360  
Qy 361 HGNDKDSVLLGNSRGSKEARASALNHNIIHSQDDKSGREIRVJHLLEOIRAYCETC 420  
Db 361 HGNDKDSVLLGNSRGSKEARASALNHNIIHSQDDKSGREIRVJHLLEOIRAYCETC 420  
Qy 421 WEMQEAHPGMDQDNVMPAPVEHQI CPAVCVLMKLSFDEEHRHAMELGLQIAELLQ 480  
Db 421 WEMQEAHPGMDQDNVMPAPVEHQI CPAVCVLMKLSFDEEHRHAMELGLQIAELLQ 480  
Qy 481 VDCENYGLTNDHYSITLARYAGMALTNLTFGDVANKATLCMKGCMRALVAQLKSEEDL 540  
Db 481 VDCENYGLTNDHYSITLARYAGMALTNLTFGDVANKATLCMKGCMRALVAQLKSEEDL 540  
Qy 541 QOVIASVLRNLSWRADVANSKTLREVGSVKALMECALEVKKESTLKSVLNLSAHT 600  
Db 541 QOVIASVLRNLSWRADVANSKTLREVGSVKALMECALEVKKESTLKSVLNLSAHT 600  
Qy 601 ENKADICAVDGALAFVLGTLTVRSQTNLTALIESGGGILRVNVSIIANEDHROILRENN 660  
Db 601 ENKADICAVDGALAFVLGTLTVRSQTNLTALIESGGGILRVNVSIIANEDHROILRENN 660  
Qy 661 CIQTLILHLKSHSLTIVSNACGTLANLSARNPKOCEALWMDGAVSMKLNLIHSKHKMIAM 720  
Db 661 CIQTLILHLKSHSLTIVSNACGTLANLSARNPKOCEALWMDGAVSMKLNLIHSKHKMIAM 720  
Qy 721 GSAALRNLMANRPARYKADANIMSPGSSILPSLHVKKOKALEAELDAQHLSLTFDNDLS 780  
Db 721 GSAALRNLMANRPARYKADANIMSPGSSILPSLHVKKOKALEAELDAQHLSLTFDNDLS 780  
Qy 781 PKASHRSKORHQSLYGDVFTNDRHDDNRSDFNFTGNMTVLSPLYNTTLPSSSSRGS 840  
Db 781 PKASHRSKORHQSLYGDVFTNDRHDDNRSDFNFTGNMTVLSPLYNTTLPSSSSRGS 840  
Qy 841 LQSSRSKDRSLERERIGLGNVHPATENPGTSSKRGLOISTTAAQIAKWMEEVSAIHTS 900  
Db 841 LQSSRSKDRSLERERIGLGNVHPATENPGTSSKRGLOISTTAAQIAKWMEEVSAIHTS 900  
Qy 901 QEDRSSGTTTELHCVTDERNALRRSAAHTHNTYNTFTKSENSNRTCSPYAKLEYKRS 960  
Db 901 QEDRSSGTTTELHCVTDERNALRRSAAHTHNTYNTFTKSENSNRTCSPYAKLEYKRS 960  
Qy 961 NDSLSVSSNDGVKGQMKPSTESYEDDESDFCSYGOYPADLAHKIHSANHMDNDGE 1020  
Db 961 NDSLSVSSNDGVKGQMKPSTESYEDDESDFCSYGOYPADLAHKIHSANHMDNDGE 1020  
Qy 1021 LDTPIVNSLYKSYDEQLNSGRQSPQNERWARPKHIEDEIKOSEQRQSRNQSTTYPVYTE 1080  
Db 1021 LDTPIVNSLYKSYDEQLNSGRQSPQNERWARPKHIEDEIKOSEQRQSRNQSTTYPVYTE 1080  
Qy 1081 STDDKHLKQPHFGQOECVPIYRSGANGETNRVCSNHNQNTYSOLCOEDDYDDKP 1140  
Db 1081 STDDKHLKQPHFGQOECVPIYRSGANGETNRVCSNHNQNTYSOLCOEDDYDDKP 1140  
Qy 1141 TNYSERYSEREEHBERPNTYIKYNEEKRHVDQPIDYSLKYATDIPSSQKQSFSSKS 1200  
Db 1141 TNYSERYSEREEHBERPNTYIKYNEEKRHVDQPIDYSLKYATDIPSSQKQSFSSKS 1200  
Qy 1201 SSQSSKTEHMSSSSENTSTPSSNAKRONLHPSSAQSRSGQPKAATCKVSSINQETIQ 1260  
Db 1201 SSQSSKTEHMSSSSENTSTPSSNAKRONLHPSSAQSRSGQPKAATCKVSSINQETIQ 1260

Qy 1261 TYCVEDTPICFRCSLSSLSAEDEIGCNQVTTQBADSSANTLQIAEIKGKIGTRSAEDPV 1320  
Db 1261 TYCVEDTPICFRCSLSSLSAEDEIGCNQVTTQBADSSANTLQIAEIKGKIGTRSAEDPV 1320  
Qy 1321 SEVP-AVSOHPRTKSRLOGSSLSSESAHKAVEPPSGAKSPSKGAQOTPKPPEHYQOE 1379  
Db 1321 SEVPVSHSTLETKSRLOGSSLSSESAHKAVEPPSGAKSPSKGAQOTPKPPEHYQOE 1380  
Qy 1380 TPLMFRCCTSVSLSLDFESRSTASSVQSPCGMVSGIISPSDLDPSPCQOTMPPPSKTP 1439  
Db 1381 TPLMFRCCTSVSLSLDFESRSTASSVQSPCGMVSGIISPSDLDPSPCQOTMPPPSKTP 1440  
Qy 1440 PPPQTAQTKREVPKNAKPTAEKRESGPKQAAVNAAVQVQVLPDADTLHLHFATESTPDG 1499  
Db 1441 PPPQTAQTKREVPKNAKPTAEKRESGPKQAAVNAAVQVQVLPDADTLHLHFATESTPDG 1500  
Qy 1500 FSCSSLSLSLDEPIQKDVLRIMPVQENDNGNETESEQPKENENQEKAEKTIIDS 1559  
Db 1501 FSCSSLSLSLDEPIQKDVLRIMPVQENDNGNETESEQPKENENQEKAEKTIIDS 1560  
Qy 1560 EKLLDDSDDDDEILEECIIISAMPTKSRGKKKPAQTASKLPPPVARKPSQLPVYKLLP 1619  
Db 1561 EKLLDDSDDDDEILEECIIISAMPTKSRGKKKPAQTASKLPPPVARKPSQLPVYKLLP 1620  
Qy 1620 SQNRLOPQKHVSFTPGDDMPRVYCVGTPIINFSTATSLDLTIESPPNELAAGEVYRGA 1679  
Db 1621 SQNRLOPQKHVSFTPGDDMPRVYCVGTPIINFSTATSLDLTIESPPNELAAGEVYRGA 1680  
Qy 1680 QSGFEKRDITPIEGRSTDEAOGKTSSTVTPBLDDNKAEBGDI LAECINSAMPKGSKH 1739  
Db 1681 QSGFEKRDITPIEGRSTDEAOGKTSSTVTPBLDDNKAEBGDI LAECINSAMPKGSKH 1740  
Qy 1740 PFRVKIMDQVQAASASSAPNKNQLDGKKKTPSPVKPIQONTYRTRVRKNADSKNL 1799  
Db 1741 PFRVKIMDQVQAASASSAPNKNQLDGKKKTPSPVKPIQONTYRTRVRKNADSKNL 1800  
Qy 1800 NAEVFSNKSQKQNLKNNSKDFNDKLPNNEDRVGSPAFDPSFHHVTPTEGTPYCFSRN 1859  
Db 1801 NAEVFSNKSQKQNLKNNSKDFNDKLPNNEDRVGSPAFDPSFHHVTPTEGTPYCFSRN 1860  
Qy 1860 DLSLSLDFDDDDVDLSREKAEALRKAENKESAKVTSHTELTNSQOSANKTQAIKQPIN 1919  
Db 1861 DLSLSLDFDDDDVDLSREKAEALRKAENKESAKVTSHTELTNSQOSANKTQAIKQPIN 1920  
Qy 1920 RGQPKPILOKQSTPPQSSKDI PDGAAATDEKLQNFALIENTPVCFSHNSLSLSDIDQEN 1979  
Db 1921 RGQPKPILOKQSTPPQSSKDI PDGAAATDEKLQNFALIENTPVCFSHNSLSLSDIDQEN 1980  
Qy 1980 NKNENEPKETEPPDSQCEPSKPOASGYAPKSFHVEDTPVCFRNSLSLSLSDIDDL 2039  
Db 1981 NKNENEPKETEPPDSQCEPSKPOASGYAPKSFHVEDTPVCFRNSLSLSLSDIDDL 2040  
Qy 2040 QECISSAMPKCKPSRLKGDNEKHSPRNMGGILGEDTLDLKQIQRDPDSEHGLSPDSENF 2099  
Db 2041 QECISSAMPKCKPSRLKGDNEKHSPRNMGGILGEDTLDLKQIQRDPDSEHGLSPDSENF 2100  
Qy 2100 DWKAIQEGANSIVSLHCAAAAACLSRQASGSDPSILSLKSGISLGSPFHLTPDQEEKPF 2159  
Db 2101 DWKAIQEGANSIVSLHCAAAAACLSRQASGSDPSILSLKSGISLGSPFHLTPDQEEKPF 2160  
Qy 2160 TSNKGPRLKPEGEKSTLETKKIESKGIKGGKVKYKSLITGKVRNSSEISGOMKQPLQA 2219  
Db 2161 TSNKGPRLKPEGEKSTLETKKIESKGIKGGKVKYKSLITGKVRNSSEISGOMKQPLQA 2220  
Qy 2220 NMPISIRGRTHHTPGVNRSSSTSPVSKGPPKLTTPASKSPSEGTATTTSPRCAKPSVK 2279  
Db 2221 NMPISIRGRTHHTPGVNRSSSTSPVSKGPPKLTTPASKSPSEGTATTTSPRCAKPSVK 2280  
Qy 2280 SELSPVARQTSQIGGSSKAPSRSGRSDTSPRPAQQLSRPIQSPGRNSISPGRNGISPP 2339  
Db 2281 SELSPVARQTSQIGGSSKAPSRSGRSDTSPRPAQQLSRPIQSPGRNSISPGRNGISPP 2340

QY 2340 NKLSQLPRTSPSTASTKSSGSGKMSYTPSGQMSQQNLTKQTLGSKNASSIPRESASK 2399  
DB 2341 NKLSQLPRTSPSTASTKSSGSGKMSYTPSGQMSQQNLTKQTLGSKNASSIPRESASK 2400  
QY 2400 GLNQMNNGANGKVKELSRMSTKSSGESDRSRPVLVROSTFIKEAPPTLRKLEES 2459  
DB 2401 GLNQMNNGANGKVKELSRMSTKSSGESDRSRPVLVROSTFIKEAPPTLRKLEES 2460  
QY 2460 ASFSLSPPSPAPTRSQATPVLSPSLPDMSLSTHSSVQAGGWRKLPNLSPTIEYND 2519  
DB 2461 ASFSLSPPSPAPTRSQATPVLSPSLPDMSLSTHSSVQAGGWRKLPNLSPTIEYND 2520  
QY 2520 GRPAKHDIARSHSESRLPINSRGTWKRHSKHSSSLPRVSTWRTGTSSSILSASSE 2579  
DB 2521 GRPAKHDIARSHSESRLPINSRGTWKRHSKHSSSLPRVSTWRTGTSSSILSASSE 2580  
QY 2580 SSEKASEDEKHVNSISGTKQSKENQVSAGTWRKIKENEFSPNTNSTQTVSSGATNGAE 2639  
DB 2581 SSEKASEDEKHVNSISGTKQSKENQVSAGTWRKIKENEFSPNTNSTQTVSSGATNGAE 2640  
QY 2640 SKTLIYQAPAVSKTEDVWRIEDCPINPRSGRSPGTGNTPPVIDSVSEKANPNIKDSKD 2699  
DB 2641 SKTLIYQAPAVSKTEDVWRIEDCPINPRSGRSPGTGNTPPVIDSVSEKANPNIKDSKD 2700  
QY 2700 NOAKQNVGNSVPMRTVGLNRLTSFIQVADPQKGTIKPGQNNPVPVSETNESPIVER 2759  
DB 2701 NOAKQNVGNSVPMRTVGLNRLTSFIQVADPQKGTIKPGQNNPVPVSETNESPIVER 2760  
QY 2760 TPFSSSSSKHSSPSGTVAARVTPFNPNPRKSSADSTGSRPSQIPTPVNNNTKKRDSK 2819  
DB 2761 TPFSSSSSKHSSPSGTVAARVTPFNPNPRKSSADSTGSRPSQIPTPVNNNTKKRDSK 2820  
QY 2820 TDSTESSGTQSPKRHSGSYLVTSV 2843  
DB 2821 TDSTESSGTQSPKRHSGSYLVTSV 2844

## RESULT 4

US-10-408-765A-1970

; Sequence 1970, Application US/10408765A

; Publication No. US20040101874A1

; GENERAL INFORMATION:

; APPLICANT: Ghosh, Soumitra S.

; APPLICANT: Fahy, Eoin D.

; APPLICANT: Zhang, Bing

; APPLICANT: Gibson, Bradford W.

; APPLICANT: Taylor, Steven W.

; APPLICANT: Glenn, Gary M.

; APPLICANT: Warnock, Dale E.

; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION

; FILE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME

; FILE REFERENCE: 660088.465

; CURRENT APPLICATION NUMBER: US/10/408,765A

; CURRENT FILING DATE: 2003-04-04

; NUMBER OF SEQ ID NOS: 3077

; SOFTWARE: PASCSEQ for Windows Version 4.0

; SEQ ID NO 1970

; LENGTH: 2843

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-408-765A-1970

Query Match

Best Local Similarity 99.7%; Score 14524; DB 16; Length 2843;

Matches 2835; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 MAAASYDQLLKQVEALKWENSLRQLEEDNSHLTKLETEASNMKEVLKQLOGSIEDAM 60  
DB 1 MAAASYDQLLKQVEALKWENSLRQLEEDNSHLTKLETEASNMKEVLKQLOGSIEDAM 60  
QY 61 ASSGQIDLLERLKLNDSSNPFQVKLRKMSLRVSGSRGSCSPVPMGSPFR 120  
DB 61 ASSGQIDLLERLKLNDSSNPFQVKLRKMSLRVSGSRGSCSPVPMGSPFR 120

QY 121 RGFVNGSRESTGYLBELEKERSLLADLDKBEKEDWYQAOLQNLTKRIDSLPTENFSL 180  
DB 121 RGFVNGSRESTGYLBELEKERSLLADLDKBEKEDWYQAOLQNLTKRIDSLPTENFSL 180  
QY 181 QTDLTRQLEYEARGIRVAMEEQLCTQDMMEKRAQRRIARIQOIEKIDILRIRQLQSQAT 240  
DB 181 QTDLTRQLEYEARGIRVAMEEQLCTQDMMEKRAQRRIARIQOIEKIDILRIRQLQSQAT 240  
QY 241 EAERSQNKHETGSHDAERQNGEGVGEINWATSGNGOGSTTRMDHETASVLSSTHSA 300  
DB 241 EAERSQNKHETGSHDAERQNGEGVGEINWATSGNGOGSTTRMDHETASVLSSTHSA 300  
QY 301 PRRLTSHLGTQVEMVYSLLSMLGTHDKDDMSRTLLAMSSSDSCISMRQSGCLPLLIQLL 360  
DB 301 PRRLTSHLGTQVEMVYSLLSMLGTHDKDDMSRTLLAMSSSDSCISMRQSGCLPLLIQLL 360  
QY 361 HGNDKDSVLLGNRSGSKARARASALHNIITHSQDDDKRGRREIRVLHLLLEQIAYCETC 420  
DB 361 HGNDKDSVLLGNRSGSKARARASALHNIITHSQDDDKRGRREIRVLHLLLEQIAYCETC 420  
QY 421 MEWQEAHEFGMDQDKNPMAPVVEHOICPAVCVLMKLSFDEHHRHAMELGGLQIAELLQ 480  
DB 421 MEWQEAHEFGMDQDKNPMAPVVEHOICPAVCVLMKLSFDEHHRHAMELGGLQIAELLQ 480  
QY 481 VDCEMYGLTNDHYSITLRYAGMALTNLTFQDVANKATLCSMKGCMRALVAQLKSEEDL 540  
DB 481 VDCEMYGLTNDHYSITLRYAGMALTNLTFQDVANKATLCSMKGCMRALVAQLKSEEDL 540  
QY 541 QOVIVASVLRLNSWADVNSKTLREVGSVKALMECALEVKKESTLKSVLALNLSAHT 600  
DB 541 QOVIVASVLRLNSWADVNSKTLREVGSVKALMECALEVKKESTLKSVLALNLSAHT 600  
QY 601 ENKADICAVDGAFLVGTLYRSOTNTLAIIESGGGILRVSSLIATNEDHRQILRENN 660  
DB 601 ENKADICAVDGAFLVGTLYRSOTNTLAIIESGGGILRVSSLIATNEDHRQILRENN 660  
QY 661 CLOTLLOHLKSHSLTIVSNACGTLNLSARNPKDOEALWDMGAVSMKNTLHSHKMIAM 720  
DB 661 CLOTLLOHLKSHSLTIVSNACGTLNLSARNPKDOEALWDMGAVSMKNTLHSHKMIAM 720  
QY 721 GSAALRNLMANRPAPKADANIMSPGSSLPSSLHVRKQKALEAEALDAQHLSETFNIDNLS 780  
DB 721 GSAALRNLMANRPAPKADANIMSPGSSLPSSLHVRKQKALEAEALDAQHLSETFNIDNLS 780  
QY 781 PKASHRSKQRHQSILYGDYVFTNHRHDDNSDNFTNNTVLSPLYLNTTVLPSSSSSRGS 840  
DB 781 PKASHRSKQRHQSILYGDYVFTNHRHDDNSDNFTNNTVLSPLYLNTTVLPSSSSSRGS 840  
QY 841 LDSRSERKDSRLERERIGLGNVHPATENPCTSSKRGLOSTTTAAQIAKYWEVSAIHTS 900  
DB 841 LDSRSERKDSRLERERIGLGNVHPATENPCTSSKRGLOSTTTAAQIAKYWEVSAIHTS 900  
QY 901 QEDRSSGSTTTELHCVTDERNALRRSSAAHTSNNTYNTKSENSNRTCSMPYAKLEYKRSS 960  
DB 901 QEDRSSGSTTTELHCVTDERNALRRSSAAHTSNNTYNTKSENSNRTCSMPYAKLEYKRSS 960  
QY 961 NDSLNSVSSNDGYGKRGOMKPSIESYSEDDSEKFCYGOYPADLAHAIHSAHNDNDGE 1020  
DB 961 NDSLNSVSSNDGYGKRGOMKPSIESYSEDDSEKFCYGOYPADLAHAIHSAHNDNDGE 1020  
QY 1021 LDTPIYSLKYSEBQLNSGRQSPSQNERWARPKHIIIEDEIKQSEQRQSRNQSTTYPVTE 1080  
DB 1021 LDTPIYSLKYSEBQLNSGRQSPSQNERWARPKHIIIEDEIKQSEQRQSRNQSTTYPVTE 1080  
QY 1081 STDDKHLKFPQHFQOQECVSPYRSRGANGSETNRVSGNHGINONVQSLOEDDYEDDKP 1140  
DB 1081 STDDKHLKFPQHFQOQECVSPYRSRGANGSETNRVSGNHGINONVQSLOEDDYEDDKP 1140  
QY 1141 TNSERYSEBEQHEEBERTPNYSIKYNEEKRVDPIDYSLKYATDIPSSQKQSFSSKS 1200  
DB 1141 TNSERYSEBEQHEEBERTPNYSIKYNEEKRVDPIDYSLKYATDIPSSQKQSFSSKS 1200

1201 SSGSSKTEHSSSENTSTPSSNAKRONQHPSAASRSGOPOKAATCKVSSINQETIQ 1260  
1201 SSGSSKTEHSSSENTSTPSSNAKRONQHPSAASRSGOPOKAATCKVSSINQETIQ 1260  
1261 TYCVEDTPICSRCSLSLSSAASDEIGCNOTTQADSAANTLOAEIKGKIGTRSAEDPV 1320  
1261 TYCVEDTPICSRCSLSLSSAASDEIGCNOTTQADSAANTLOAEIKGKIGTRSAEDPV 1320  
1321 SEVPAVSQHPRTKSRLOGSSLSSESARHKAVERPSSGAKSPKSGAOTPKSPPHYVOET 1380  
1321 SEVPAVSQHPRTKSRLOGSSLSSESARHKAVERPSSGAKSPKSGAOTPKSPPHYVOET 1380  
1381 PLMFSRCTSVSSLSDFERSSTASSVQSEPCSGMWVGIIISPSDLDPSPGQTMPPRSKTPP 1440  
1381 PLMFSRCTSVSSLSDFERSSTASSVQSEPCSGMWVGIIISPSDLDPSPGQTMPPRSKTPP 1440  
1441 PPPQTAQTKREVPKNAKPTAKRESGPKQAAVNAAVQVQVLPDADTLLHFAESTPDGF 1500  
1441 PPPQTAQTKREVPKNAKPTAKRESGPKQAAVNAAVQVQVLPDADTLLHFAESTPDGF 1500  
1501 SCSSLSALSALDEPTIQKDELIRMPVQENDNGNETESEOPEKESNENOEKAEKTIIDSE 1560  
1501 SCSSLSALSALDEPTIQKDELIRMPVQENDNGNETESEOPEKESNENOEKAEKTIIDSE 1560  
1561 KDLLDDSDDDDIIELEECIIISAMPTKSSRKGKPAQATASKLPPPVVARKPSQLPVYKLLPS 1620  
1561 KDLLDDSDDDDIIELEECIIISAMPTKSSRKGKPAQATASKLPPPVVARKPSQLPVYKLLPS 1620  
1621 QNRLOPQKHVSFTPODDMPRVYCEGTPIINSTATSLSDLIIESPPNELAAGEVGRGAQ 1680  
1621 QNRLOPQKHVSFTPODDMPRVYCEGTPIINSTATSLSDLIIESPPNELAAGEVGRGAQ 1680  
1681 SGPEFKRTIPTEGRSTDEAOGKTSVTIPELDNKAEEGDIIAECINSAMPKKGSHK 1740  
1681 SGPEFKRTIPTEGRSTDEAOGKTSVTIPELDNKAEEGDIIAECINSAMPKKGSHK 1740  
1741 FRVKKIMQVQOASASSAPNKNQLDGKKKPTSPVKDIPONTYRTRVRKNASKNLN 1800  
1741 FRVKKIMQVQOASASSAPNKNQLDGKKKPTSPVKDIPONTYRTRVRKNASKNLN 1800  
1801 ABRVFSNDKSKONLKNNSKDFNDKLNEDRVGSGFAFSPHYTPIEGTPVCFERN 1860  
1801 ABRVFSNDKSKONLKNNSKDFNDKLNEDRVGSGFAFSPHYTPIEGTPVCFERN 1860  
1861 SLSSLDFFDDDDVLDGREKAELEKAKENKESAKVTSHTELSNQOASANKTQAIKQPINR 1920  
1861 SLSSLDFFDDDDVLDGREKAELEKAKENKESAKVTSHTELSNQOASANKTQAIKQPINR 1920  
1921 GQPKILOQSTFPOSSKDIIPRGAATDEKLNFAIENTPVCFSHNSLSLSIDQEN 1980  
1921 GQPKILOQSTFPOSSKDIIPRGAATDEKLNFAIENTPVCFSHNSLSLSIDQEN 1980  
1981 NKENEPKETEPPDQSGEPSPKQASGYAPKSPHVEDTPVCFSRNSSLSSLSIDGEDLLQ 2040  
1981 NKENEPKETEPPDQSGEPSPKQASGYAPKSPHVEDTPVCFSRNSSLSSLSIDGEDLLQ 2040  
2041 ECISAMPKPKKPSRLKGNEXHSPRNMGGIIGEDTLDLKDIQRPDSEHGLSPDSENF 2100  
2041 ECISAMPKPKKPSRLKGNEXHSPRNMGGIIGEDTLDLKDIQRPDSEHGLSPDSENF 2100  
2101 WKAIQEGANSIYVSLHQAAAACLRQASSDSDSLSLKSGISLSPHLPDQEKFT 2160  
2101 WKAIQEGANSIYVSLHQAAAACLRQASSDSDSLSLKSGISLSPHLPDQEKFT 2160  
2161 SNKGRPILKPGKSTLETKKIESEKIGKGGKVKYKSLITGVRNSNIEISGOMKQLOAN 2220  
2161 SNKGRPILKPGKSTLETKKIESEKIGKGGKVKYKSLITGVRNSNIEISGOMKQLOAN 2220  
2221 MFSISGRMTIHPGVNRSSSTSPVSKGPKLPTPKASPSSEGTATTSRPGAKPSVKS 2280  
2221 MFSISGRMTIHPGVNRSSSTSPVSKGPKLPTPKASPSSEGTATTSRPGAKPSVKS 2280  
2281 ELSPVARQTSQIGGSKAPSRSGRSDTSPRPAQOPLSRPIOSPGRNSISPGRNGISPPN 2340

2281 ELSPVARQTSQIGGSKAPSRSGRSDTSPRPAQOPLSRPIOSPGRNSISPGRNGISPPN 2340  
2341 KLSQLPRTSPSTASTKSSGSKMSYTSRQMSQONLTQOTGLSKNAGSIIPRSESASG 2400  
2341 KLSQLPRTSPSTASTKSSGSKMSYTSRQMSQONLTQOTGLSKNAGSIIPRSESASG 2400  
2401 LNMNNGANGANKVELSRMSSTKSGSESDRERPVLRQSTFIKEAPSTLRRKLEESA 2460  
2401 LNMNNGANGANKVELSRMSSTKSGSESDRERPVLRQSTFIKEAPSTLRRKLEESA 2460  
2461 SPESLSPSRPASPTRSQAOTPVLSPLDMSLSTHSSVQAGGWRKLPNLSPTIEYNDG 2520  
2461 SPESLSPSRPASPTRSQAOTPVLSPLDMSLSTHSSVQAGGWRKLPNLSPTIEYNDG 2520  
2521 RPAKRHDIAARSHSESRLPINESGTWKEHSHSSSLPRVSTWRTGSSSSTLSASSS 2580  
2521 RPAKRHDIAARSHSESRLPINESGTWKEHSHSSSLPRVSTWRTGSSSSTLSASSS 2580  
2581 SEKAKSEDEKHVNSISGTQSKENQVSAKGTWRKIKENEFPSTNSTQTVSSGATNGAES 2640  
2581 SEKAKSEDEKHVNSISGTQSKENQVSAKGTWRKIKENEFPSTNSTQTVSSGATNGAES 2640  
2641 KTLIYOMAPAVSKTEDVWRIEDCPINNPRSGSPGTGNTPPVIDSVSEKANPNIKDKON 2700  
2641 KTLIYOMAPAVSKTEDVWRIEDCPINNPRSGSPGTGNTPPVIDSVSEKANPNIKDKON 2700  
2701 QAKQNVGNSVPMRTVGLNRLTSFIQVADPOKGTETIKPGQNNPVVSETNESPIVERT 2760  
2701 QAKQNVGNSVPMRTVGLNRLTSFIQVADPOKGTETIKPGQNNPVVSETNESPIVERT 2760  
2761 PFSSSSSSKSHSPSGTVAARVTFNPNPRKSSADSTARPSPQIPTPVNNNTKKRDSKT 2820  
2761 PFSSSSSSKSHSPSGTVAARVTFNPNPRKSSADSTARPSPQIPTPVNNNTKKRDSKT 2820  
2821 DSTESSGTOQPKHSGSYLWTSV 2843  
2821 DSTESSGTOQPKHSGSYLWTSV 2843

## RESULT 5

US-08-681-219-32  
; Sequence 32, Application US/08681219  
; Publication NO. US20020058607A1  
; GENERAL INFORMATION:  
; APPLICANT: Takaaki Sato and Junn Yanagisawa  
; TITLE OF INVENTION: COMPOUNDS THAT INHIBIT THE INTERACTION BETWEEN  
; TITLE OF INVENTION: SIGNAL-TRANSDUCING PROTEINS AND THE GLGF  
; TITLE OF INVENTION: (PDZ/DHR) DOMAIN AND USES THEREOF  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper & Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/681,219  
; FILING DATE: 22-JUL-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 0575/48962/JPW/JKM  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 278-0400  
; TELEFAX: (212) 391-0525

; INFORMATION FOR SEQ ID NO: 32:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 2843 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 US-08-681-219-32

Query Match 99.5%; Score 14506; DB 8; Length 2843;  
 Best Local Similarity 99.5%; Pred. No. 0;  
 Matches 2830; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY	1	MAAASYDQLLKQV	ALAEVSNL	ROELEDNS	NHLLTK	LETEASN	KEVILK	QGSIDEAM	60
DB	1	MAAASYDQLLKQV	ALAEVSNL	RQELLEDNS	NHLLTK	LETEASN	KEVILK	QGSIDEAM	60
QY	61	ASSGQIDLLERL	KEINLDSN	FPGVKLR	SKMSLR	SYSGREG	SVSRSG	SGCSPVPMGSFPR	120
DB	61	ASSGQIDLLERL	KEINLDSN	FPGVKLR	SKMSLR	SYSGREG	SVSRSG	SGCSPVPMGSFPR	120
QY	121	RGFVNGSR	ESTGYLE	EELEKERS	LLADL	DKBEKE	KWYAAQ	LNLTKR	DSLPLTENFSL
DB	121	RGFVNGSR	ESTGYLE	EELEKERS	LLADL	DKBEKE	KWYAAQ	LNLTKR	DSLPLTENFSL
QY	181	QTDLTR	ROLEYEAR	QIRVAME	EOELGT	QDMKRA	QRRIARI	QIQIEKD	ILIRIOLLOSOAT
DB	181	QTDLTR	ROLEYEAR	QIRVAME	EOELGT	QDMKRA	QRRIARI	QIQIEKD	ILIRIOLLOSOAT
QY	241	EAERSSQNK	HETGSH	DAERNEG	GVGE	INATSG	NGQGST	TRMDHETAS	VLSSSSSTHSA
DB	241	EAERSSQNK	HETGSH	DAERNEG	GVGE	INATSG	NGQGST	TRMDHETAS	VLSSSSSTHSA
QY	301	PRRLTSL	GTKEVMY	SLLMGL	THDKDD	MSRTLL	AMSSOD	SCISWROSG	CLPLLIQLL
DB	301	PRRLTSL	GTKEVMY	SLLMGL	THDKDD	MSRTLL	AMSSOD	SCISWROSG	CLPLLIQLL
QY	361	HGNDKDS	VLLGNR	SGKEAR	ASAAAL	HNIIHS	QPDQK	RREIRV	LHLEQIRAYCETC
DB	361	HGNDKDS	VLLGNR	SGKEAR	ASAAAL	HNIIHS	QPDQK	RREIRV	LHLEQIRAYCETC
QY	421	WEQEA	HEPGMDQ	KNPMPAP	VEHQIC	PAVCV	LMKLS	FDEEHR	HANVELGGLOAIELLO
DB	421	WEQEA	HEPGMDQ	KNPMPAP	VEHQIC	PAVCV	LMKLS	FDEEHR	HANVELGGLOAIELLO
QY	481	VDCEMY	GLTNDH	YSITLRR	YAGMAL	TNLTP	GDVANK	ATLCS	MKGCMALVAQLKSESDI
DB	481	VDCEMY	GLTNDH	YSITLRR	YAGMAL	TNLTP	GDVANK	ATLCS	MKGCMALVAQLKSESDI
QY	541	QQVIAS	VLRLN	SWRAD	NSKTTREV	GSVKAL	MECAL	EVKES	TLSKSVLSALWLSAHT
DB	541	QQVIAS	VLRLN	SWRAD	NSKTTREV	GSVKAL	MECAL	EVKES	TLSKSVLSALWLSAHT
QY	601	ENKADI	CAVDG	ALAFV	GLTYRS	QNTLAI	IBSGG	GILRN	VSSLIATNEDH
DB	601	ENKADI	CAVDG	ALAFV	GLTYRS	QNTLAI	IBSGG	GILRN	VSSLIATNEDH
QY	661	CLQTL	LOHLKSH	LTVSN	ACGLT	WNLSAR	PNKDP	QALMD	WAGVMSLKNLIHSKHMIA
DB	661	CLQTL	LOHLKSH	LTVSN	ACGLT	WNLSAR	PNKDP	QALMD	WAGVMSLKNLIHSKHMIA
QY	721	GSAAL	ENLW	ANP	AKYK	DIANS	PGSSLP	SLHVR	KQKALELDAQHLSETF
DB	721	GSAAL	ENLW	ANP	AKYK	DIANS	PGSSLP	SLHVR	KQKALELDAQHLSETF
QY	781	PKASH	RSKR	QKQSLY	CDYV	FTNR	HDNR	SNFN	TGNMTVLSPLYNTITVLPSSSSSRGS
DB	781	PKASH	RSKR	QKQSLY	CDYV	FTNR	HDNR	SNFN	TGNMTVLSPLYNTITVLPSSSSSRGS
QY	841	LDSSR	SEKORS	LERE	RIGIG	GNVH	PATEN	PGTSS	KEGLQISTTAAQIAKMEVSAIHTS
DB	841	LDSSR	SEKORS	LERE	RIGIG	GNVH	PATEN	PGTSS	KEGLQISTTAAQIAKMEVSAIHTS

QY	901	QEDRSSG	STTELHC	VTDER	NALRR	SSAAHT	SNNTY	NFTK	SENSNR	TC	SM	PYAK	LEYKR	SS	960
DB	901	QEDRSSG	STTELHC	VTDER	NALRR	SSAAHT	SNNTY	NFTK	SENSNR	TC	SM	PYAK	LEYKR	SS	960
QY	961	NDLS	NSVSN	DYK	GK	GOMK	PSIES	YSED	DESK	CSY	GQY	PAD	LAH	KI	THS
DB	961	NDLS	NSVSN	DYK	GK	GOMK	PSIES	YSED	DESK	CSY	GQY	PAD	LAH	KI	THS
QY	1021	LDTP	INYS	GLK	YSD	EQ	LN	SGR	QSP	QNE	RW	APK	HI	IE	DEI
DB	1021	LDTP	INYS	GLK	YSD	EQ	LN	SGR	QSP	QNE	RW	APK	HI	IE	DEI
QY	1081	STDD	KHLK	FQ	HC	QO	ECV	SPY	R	SG	ANG	SET	NR	V	GN
DB	1081	STDD	KHLK	FQ	HC	QO	ECV	SPY	R	SG	ANG	SET	NR	V	GN
QY	1141	TNYS	ERY	EEEE	QH	EEEE	EP	TNYS	IKY	NE	KH	V	DP	IDY	SL
DB	1141	TNYS	ERY	EEEE	QH	EEEE	EP	TNYS	IKY	NE	KH	V	DP	IDY	SL
QY	1201	SSGQ	SKT	TEH	MS	SS	SENT	STP	SS	NA	KQ	NQ	LHP	SS	AQ
DB	1201	SSGQ	SKT	TEH	MS	SS	SENT	STP	SS	NA	KQ	NQ	LHP	SS	AQ
QY	1261	TYCV	EDT	PI	CF	SR	CS	SL	SS	LA	ES	AE	IG	CM	TO
DB	1261	TYCV	EDT	PI	CF	SR	CS	SL	SS	LA	ES	AE	IG	CM	TO
QY	1321	SEV	PAV	SQ	HP	TK	SR	SL	SS	LA	ES	AE	IG	CM	TO
DB	1321	SEV	PAV	SQ	HP	TK	SR	SL	SS	LA	ES	AE	IG	CM	TO
QY	1381	PLMF	SR	CT	SV	SL	DS	FE	SR	ST	ASS	VQ	SE	PC	SG
DB	1381	PLMF	SR	CT	SV	SL	DS	FE	SR	ST	ASS	VQ	SE	PC	SG
QY	1441	PPPT	QA	TK	RE	VP	KN	KA	PT	AE	K	RES	GP	QA	AA
DB	1441	PPPT	QA	TK	RE	VP	KN	KA	PT	AE	K	RES	GP	QA	AA
QY	1501	SCSS	SL	SA	LS	LD	EF	IO	KD	VEL	R	IMP	VO	EN	D
DB	1501	SCSS	SL	SA	LS	LD	EF	IO	KD	VEL	R	IMP	VO	EN	D
QY	1561	KD	LD	DD	DD	DD	DD	DD	DD	DD	DD	DD	DD	DD	DD
DB	1561	KD	LD	DD	DD	DD	DD	DD	DD	DD	DD	DD	DD	DD	DD
QY	1621	QNR	LO	P	OK	H	S	F	T	PG	D	M	P	R	V
DB	1621	QNR	LO	P	OK	H	S	F	T	PG	D	M	P	R	V
QY	1681	SGE	PE	K	D	T	I	P	T	E	R	S	T	E	A
DB	1681	SGE	PE	K	D	T	I	P	T	E	R	S	T	E	A
QY	1741	FRV	K	I	M	O	V	Q	A	S	A	S	A	S	A
DB	1741	FRV	K	I	M	O	V	Q	A	S	A	S	A	S	A
QY	1801	AER	V	S	N	K	S	K	N	L	N	S	K	D	F
DB	1801	AER	V	S	N	K	S	K	N	L	N	S	K	D	F
QY	1861	SL	S	L	D	D	D	D	D	D	D	D	D	D	D
DB	1861	SL	S	L	D	D	D	D	D	D	D	D	D	D	D
QY	1921	GQ	P	I	L	O	K	O	S	T	P	O	S	S	K
DB	1921	GQ	P	I	L	O	K	O	S	T	P	O	S	S	K
QY	1981	NK	NE	P	I	K	E	T	E	P	P	D	S	G	
DB	1981	NK	NE	P	I	K	E	T	E	P	P	D	S	G	



1981 NKENPIKETEPDPSQEPKQASGVAPKSFHVEDTPVCFERNSSLSLSISEDDLLQ 2040  
2041 ECISAMPKKKPKRLKGDNEKHSRPNVGGILGEDTLDKDQIOPDSEHGLSPDSNFD 2100  
2041 ECISAMPKKKPKRLKGDNEKHSRPNVGGILGEDTLDKDQIOPDSEHGLSPDSNFD 2100  
2101 WKATOEGANSTVSSIHQAAAAACLSROASDSDSILSKSGISLGSFPHLTPDQEEKPF 2160  
2101 WKATOEGANSTVSSIHQAAAAACLSROASDSDSILSKSGISLGSFPHLTPDQEEKPF 2160  
2161 SNKGPRILKPEKSTLETETKTESKGIKGGKVKYKSLITKVRNSSEISGQMKQIOLAN 2220  
2161 SNKGPRILKPEKSTLETETKTESKGIKGGKVKYKSLITKVRNSSEISGQMKQIOLAN 2220  
2221 MPSISRGRTMTIHIGVRNSSSTSPVSKKGPPLKTPASKSPSEGTATTSPRGAKPSVKS 2280  
2221 MPSISRGRTMTIHIGVRNSSSTSPVSKKGPPLKTPASKSPSEGTATTSPRGAKPSVKS 2280  
2281 ELSPVARTSQIIGSSKAPSGSRDSTPPSPAQPLSRPIQSPGRNSISPRNGISPPN 2340  
2281 ELSPVARTSQIIGSSKAPSGSRDSTPPSPAQPLSRPIQSPGRNSISPRNGISPPN 2340  
2341 KLSQLPRTSSPTASTKSGSGKMSYTSPPGRMQQNLTKQTGLSKVASSIPRSESASG 2400  
2341 KLSQLPRTSSPTASTKSGSGKMSYTSPPGRMQQNLTKQTGLSKVASSIPRSESASG 2400  
2401 LNMNNGANKVVELSRMSTKSGSDSRERPLVROQTFIKEAPSPFLRRKLBESA 2460  
2401 LNMNNGANKVVELSRMSTKSGSDSRERPLVROQTFIKEAPSPFLRRKLBESA 2460  
2461 SFESLSPSRPASPTRSQAQTPVLSPLDMSLTHSSVQAGGWRKLPPLNLSPTIEYNDG 2520  
2461 SFESLSPSRPASPTRSQAQTPVLSPLDMSLTHSSVQAGGWRKLPPLNLSPTIEYNDG 2520  
2521 RPAKHDIARSHESPSRLPTNRESGTWRKSHSSSLPRVSTWERTGSSSSILSASSES 2580  
2521 RPAKHDIARSHESPSRLPTNRESGTWRKSHSSSLPRVSTWERTGSSSSILSASSES 2580  
2581 SEKAKSEDEKHVNSTSGTKQKENVSAKGTWRKIKENEFSPTNSTSTVSSGATNGAES 2640  
2581 SEKAKSEDEKHVNSTSGTKQKENVSAKGTWRKIKENEFSPTNSTSTVSSGATNGAES 2640  
2641 KTLIIQMAPAVSKTDDVVRVIEDCPINNPRSGRPTGNTPTVIDSVSEKAPNIKSDKN 2700  
2641 KTLIIQMAPAVSKTDDVVRVIEDCPINNPRSGRPTGNTPTVIDSVSEKAPNIKSDKN 2700  
2701 QAKQVNGSVPMRTVGLNRLTSTIQVADPQKTEIKPGONNPVPVSEINNEPIVERT 2760  
2701 QAKQVNGSVPMRTVGLNRLTSTIQVADPQKTEIKPGONNPVPVSEINNEPIVERT 2760  
2761 PFSSSSSKHSFGTVAARVTPFNYPNPRKSSADTSARPSQIPTPVNNNTKRDST 2820  
2761 PFSSSSSKHSFGTVAARVTPFNYPNPRKSSADTSARPSQIPTPVNNNTKRDST 2820  
2821 DSTESSGTOSPKRHSGLVTSV 2843  
2821 DSTESSGTOSPKRHSGLVTSV 2843

RESULT 6  
US-09-230-111C-30  
; Sequence 30, Application US/09230111C  
; Publication No. US20030203414A1  
; GENERAL INFORMATION:  
; APPLICANT: Sato, Taka-Aki  
; TITLE OF INVENTION: COMPOUNDS THAT INHIBIT INTERACTION BETWEEN  
; TITLE OF INVENTION: SIGNAL-TRANSDUCING PROTEINS AND THE GLCF (PDZ/DHR)  
; TITLE OF INVENTION: DOMAIN AND USES THEREOF  
; FILE REFERENCE: 48962-A-PCT-US  
; CURRENT APPLICATION NUMBER: US/09/230,111C  
; CURRENT FILING DATE: 1999-05-17

NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 30  
; LENGTH: 2843  
; TYPE: PRT  
; ORGANISM: human  
US-09-230-111C-30  
Query Match 99.5%; Score 14506; DB 11; Length 2843;  
Best Local Similarity 99.5%; Pred. No. 0;  
Matches 2830; Conservative 5; Mismatches 8; Indels 0; Gaps 0;  
QY 1 MAAASYDQLLKQVEALKMENSNLROELEDNSNHLTKLETEASNMKVLEVKLOGSIDEAM 60  
Db 1 MAAASYDQLLKQVEALKMENSNLROELEDNSNHLTKLETEASNMKVLEVKLOGSIDEAM 60  
QY 61 ASSGQIDLLERLKEINLNDSSNPPGVKLRKMSLRSYGSRGVSRSRSGSCSPVPMGSFPR 120  
Db 61 ASSGQIDLLERLKEINLNDSSNPPGVKLRKMSLRSYGSRGVSRSRSGSCSPVPMGSFPR 120  
QY 121 RGFVNGSRESTGYLELEKERSILLADLDKEKEKDWYVAQNLTKRIDSPLTFNFSL 180  
Db 121 RGFVNGSRESTGYLELEKERSILLADLDKEKEKDWYVAQNLTKRIDSPLTFNFSL 180  
QY 181 QDLTRQLEYEARQIRVAMEBOLGTCQDMEKRAQRRIARIOQIEKDIIRIRQLLOSQAT 240  
Db 181 QDLTRQLEYEARQIRVAMEBOLGTCQDMEKRAQRRIARIOQIEKDIIRIRQLLOSQAT 240  
QY 241 EADRSQNHETGSHDAERQNEQGVGEINMATSNGQSGSTTRMDHETASVLSSSSTHSA 300  
Db 241 EADRSQNHETGSHDAERQNEQGVGEINMATSNGQSGSTTRMDHETASVLSSSSTHSA 300  
QY 301 PRRLTSHLGTVMYVSLLSMLGTHDKDDMSRTLLAMSSSDSCISMRQSGCLPLLIQLL 360  
Db 301 PRRLTSHLGTVMYVSLLSMLGTHDKDDMSRTLLAMSSSDSCISMRQSGCLPLLIQLL 360  
QY 361 HGMNDKSVLLNGSRGSEKARASAAHNIHSDPDKGRREIRVHLHLEQIRAYCETC 420  
Db 361 HGMNDKSVLLNGSRGSEKARASAAHNIHSDPDKGRREIRVHLHLEQIRAYCETC 420  
QY 421 WEQEAEHPGMDQKMPMPAPVEHQICPAVCVLMKLSFDEEHRHANNELGGLQAIALLQ 480  
Db 421 WEQEAEHPGMDQKMPMPAPVEHQICPAVCVLMKLSFDEEHRHANNELGGLQAIALLQ 480  
QY 481 VDCEMYGLTNDHYSITLRRYAGWALNTLTFDGVANKATLCSMGKCMRALVAQLKSSSEDL 540  
Db 481 VDCEMYGLTNDHYSITLRRYAGWALNTLTFDGVANKATLCSMGKCMRALVAQLKSSSEDL 540  
QY 541 QQVIASVLRNLSWRADVNSKKTIREVGSVKALMECALEVKKESTLKSVALNLSAHC 600  
Db 541 QQVIASVLRNLSWRADVNSKKTIREVGSVKALMECALEVKKESTLKSVALNLSAHC 600  
QY 601 ENKADICAVDGAFLVGLTYRSQNTLAIIESGGILRNVSLLIATNEDHRCILRENN 660  
Db 601 ENKADICAVDGAFLVGLTYRSQNTLAIIESGGILRNVSLLIATNEDHRCILRENN 660  
QY 661 CLQTLQHLKSHSLTIVSNACGTLNLSARNPKDQALWDMGAVSMLKNIHSGKHMAM 720  
Db 661 CLQTLQHLKSHSLTIVSNACGTLNLSARNPKDQALWDMGAVSMLKNIHSGKHMAM 720  
QY 721 GSAALNLMANRPAPKYDANIMSPGSSLSLHVRRKOKALEAELDAQHSETFDNINLS 780  
Db 721 GSAALNLMANRPAPKYDANIMSPGSSLSLHVRRKOKALEAELDAQHSETFDNINLS 780  
QY 781 PKASHRSKQHKQSLYDGYVFDNTRHDDNRSDFNTGNMTVLPYLNNTVLPSSSSSRGS 840  
Db 781 PKASHRSKQHKQSLYDGYVFDNTRHDDNRSDFNTGNMTVLPYLNNTVLPSSSSSRGS 840  
QY 841 LDSSRSKDRSLERERIGLGNVHPATENPGTSSKRGLOISTTAAQIAKVMEEVSAIHTS 900  
Db 841 LDSSRSKDRSLERERIGLGNVHPATENPGTSSKRGLOISTTAAQIAKVMEEVSAIHTS 900  
QY 901 QEDRSSGSTTELHCVTDERNALRESSAAHTSNITNTKSENSNRTCSMPYAKLEYKRSS 960





; LENGTH: 2843			
; TYPE: PRT			
; ORGANISM: human			
US-10-092-138-30			
Query Match			
Best Local Similarity 99.5%; Score 14506; DB 14; Length 2843;			
Pred. No. 0;			
Matches 2830; Conservative 5; Mismatches 8; Indels 0; Gaps 0;			
Qy	1	MAAASYDQLLKQVEALKNENLNRLQELDNENHITKLETEASNMKEVLKOLGSIIEDEAM	60
Db	1	MAAASYDQLLKQVEALKNENLNRLQELDNENHITKLETEASNMKEVLKOLGSIIEDEAM	60
Qy	61	ASSGQIDLLERLKEINLIDSSNFPFGVKLRKMSLSYSGRSREGSVSRSGECSPVPMGSPFR	120
Db	61	ASSGQIDLLERLKEINLIDSSNFPFGVKLRKMSLSYSGRSREGSVSRSGECSPVPMGSPFR	120
Qy	121	RGFVNGSRETCYELERKESLLIADLDKEKEDWYAOQLNLTFRIDSLPTENFSL	180
Db	121	RGFVNGSRETCYELERKESLLIADLDKEKEDWYAOQLNLTFRIDSLPTENFSL	180
Qy	181	QTDLTRRQLEVEARQIRVAMEBQGLTQCDMEKRAQRRIARIQEKLIRIRQLLQSOAT	240
Db	181	QTDLTRRQLEVEARQIRVAMEBQGLTQCDMEKRAQRRIARIQEKLIRIRQLLQSOAT	240
Qy	241	EAERSQNKHETGSHDARQNEGGQVGEINMATSGQGQSTTRMDHETASVLSSTHSA	300
Db	241	EAERSQNKHETGSHDARQNEGGQVGEINMATSGQGQSTTRMDHETASVLSSTHSA	300
Qy	301	PRRLTSHLGTKEVYVYSLMLGTHDKDDMSRTLLAMSSQDSCTISVQSGCLPLLIQLL	360
Db	301	PRRLTSHLGTKEVYVYSLMLGTHDKDDMSRTLLAMSSQDSCTISVQSGCLPLLIQLL	360
Qy	361	HGNDKXSVLLGNRSGSKARAPASAAALHNIHSQPDCKRGRREIRVLLHLLQIRAYCETC	420
Db	361	HGNDKXSVLLGNRSGSKARAPASAAALHNIHSQPDCKRGRREIRVLLHLLQIRAYCETC	420
Qy	421	NEWQBAHPGMDQXNPAPVEHOICPAVCVLMKLSFDEHRHNMNELGLQIAELQL	480
Db	421	NEWQBAHPGMDQXNPAPVEHOICPAVCVLMKLSFDEHRHNMNELGLQIAELQL	480
Qy	481	VDCEMYGLTNDHYSITLRYAGMALTNLTFFGVANKATLCMKGCMRALVAQLKSEBDL	540
Db	481	VDCEMYGLTNDHYSITLRYAGMALTNLTFFGVANKATLCMKGCMRALVAQLKSEBDL	540
Qy	541	QOVIASVLRNLSWRADVNSKTLREVGSVKALMECALEVYKKESTLKSVALNLSAHCT	600
Db	541	QOVIASVLRNLSWRADVNSKTLREVGSVKALMECALEVYKKESTLKSVALNLSAHCT	600
Qy	601	ENKADI CAVDGALAFVLGTLTYRSQNTLAIIESGGIILNVVSLIATNEDHRQLRENN	660
Db	601	ENKADI CAVDGALAFVLGTLTYRSQNTLAIIESGGIILNVVSLIATNEDHRQLRENN	660
Qy	661	CLQTLLOHLKSHSLTIVSNACGTLNLSARNPKDQEALWDMGAVSMKLNLIHSHKMIAM	720
Db	661	CLQTLLOHLKSHSLTIVSNACGTLNLSARNPKDQEALWDMGAVSMKLNLIHSHKMIAM	720
Qy	721	GSAALRNLMNRPAPYKXDNIMSPGSSLP SLHVKKOKALEAELDAOHLSETFDNIDLS	780
Db	721	GSAALRNLMNRPAPYKXDNIMSPGSSLP SLHVKKOKALEAELDAOHLSETFDNIDLS	780
Qy	781	PKASHRSKORHKOSLYGDVYFTNRRHDDNRSDFNTGNMTVLSPLYNTITVLPSSSSRGS	840
Db	781	PKASHRSKORHKOSLYGDVYFTNRRHDDNRSDFNTGNMTVLSPLYNTITVLPSSSSRGS	840
Qy	841	LSSRSSEKDRSLERERIGLIGNYHPATENPTGSSKRGLOISITAAQIAKMWEEVSAIHTS	900
Db	841	LSSRSSEKDRSLERERIGLIGNYHPATENPTGSSKRGLOISITAAQIAKMWEEVSAIHTS	900
Qy	901	QEDRSSGSGTTELHCVTDERNALLRRSSAAHTNTYNTFTKSENRTCSMPYAKLYKXSS	960
Db	901	QEDRSSGSGTTELHCVTDERNALLRRSSAAHTNTYNTFTKSENRTCSMPYAKLYKXSS	960

Qy	961	NDLSNVSSNDGYGKRGQMKPSIESYSEDDSEKFCYGYQYPADLAHKIHSANHMDNDGE	1020
Db	961	NDLSNVSSNDGYGKRGQMKPSIESYSEDDSEKFCYGYQYPADLAHKIHSANHMDNDGE	1020
Qy	1021	LDTPIYNSLYKSYDEQLNSGRQSPQNERWARPKHIIIEDEIKOSEQOSRQNSTTTPVYITE	1080
Db	1021	LDTPIYNSLYKSYDEQLNSGRQSPQNERWARPKHIIIEDEIKOSEQOSRQNSTTTPVYITE	1080
Qy	1081	STDDKHLKQPHFGQOECVSPYRSRGANGSEINRVGSHNGINQNVQSLSLQCEDDYEDDKP	1140
Db	1081	STDDKHLKQPHFGQOECVSPYRSRGANGSEINRVGSHNGINQNVQSLSLQCEDDYEDDKP	1140
Qy	1141	TNYSERYSEEEHEERPTNYSIKYNEBKRVHDQIDYSLKYATNDIPSSQKQSPFSKSS	1200
Db	1141	TNYSERYSEEEHEERPTNYSIKYNEBKRVHDQIDYSLKYATNDIPSSQKQSPFSKSS	1200
Qy	1201	SSQSSKTEHMSSSSNTSTPSSNAKRONQLPSSAQSSGQPOKAAATCKVSINQETIQ	1260
Db	1201	SSQSSKTEHMSSSSNTSTPSSNAKRONQLPSSAQSSGQPOKAAATCKVSINQETIQ	1260
Qy	1261	TYCVETDTPICFSCSSLSLSAEDIGCQNTTQEADSAANTLQIAIKIGKIGTRSAEDPV	1320
Db	1261	TYCVETDTPICFSCSSLSLSAEDIGCQNTTQEADSAANTLQIAIKIGKIGTRSAEDPV	1320
Qy	1321	SEYPAVSOHPRTKSSRLQSSLSSESARHKAVEFPSSGAKSPSKSGAQTPKSPPEHYVOET	1380
Db	1321	SEYPAVSOHPRTKSSRLQSSLSSESARHKAVEFPSSGAKSPSKSGAQTPKSPPEHYVOET	1380
Qy	1381	PLMFSRCTSVSSLDSPESRSIASSVQSEPCSGMVSGIISPSDLPSPGQTMPPSRSKTTP	1440
Db	1381	PLMFSRCTSVSSLDSPESRSIASSVQSEPCSGMVSGIISPSDLPSPGQTMPPSRSKTTP	1440
Qy	1441	PPQTAQTKREVPKNKAPTAEKRESGPKQAANVAARVQVLPDADTLHLFAFESTPDGF	1500
Db	1441	PPQTAQTKREVPKNKAPTAEKRESGPKQAANVAARVQVLPDADTLHLFAFESTPDGF	1500
Qy	1501	SCSSLSALSLDSPFFIQKVELRIMPVQENDNGNETSEBPKESENEQKEAEKIDSE	1560
Db	1501	SCSSLSALSLDSPFFIQKVELRIMPVQENDNGNETSEBPKESENEQKEAEKIDSE	1560
Qy	1561	KDLDDSDDDDDIIELEECIISAMPTKSSRKGPQAOTASKLPPPVARKPSQLPVYKLLPS	1620
Db	1561	KDLDDSDDDDDIIELEECIISAMPTKSSRKGPQAOTASKLPPPVARKPSQLPVYKLLPS	1620
Qy	1621	QNRLOPKHVSFTPGDDMPRVYCVETGTPINFSTATSLDLTIESPPNELAAGVGVRGAQ	1680
Db	1621	QNRLOPKHVSFTPGDDMPRVYCVETGTPINFSTATSLDLTIESPPNELAAGVGVRGAQ	1680
Qy	1681	SGEFKEDTTIPTEGRSTDEAQQGKTSSVTIPELDDNKAEGDILABCINSAMPKGSHPK	1740
Db	1681	SGEFKEDTTIPTEGRSTDEAQQGKTSSVTIPELDDNKAEGDILABCINSAMPKGSHPK	1740
Qy	1741	FRVKKIMDQVOQASASSAPNKNQDCKKKKPTSPVKPIQNTTEYRTRVKNADSNNLN	1800
Db	1741	FRVKKIMDQVOQASASSAPNKNQDCKKKKPTSPVKPIQNTTEYRTRVKNADSNNLN	1800
Qy	1801	AERFSDNCKSKONLKNNSKDFNDKLPNNEDRVGRSFAFDSPHHYTPIEGTFCFSRND	1860
Db	1801	AERFSDNCKSKONLKNNSKDFNDKLPNNEDRVGRSFAFDSPHHYTPIEGTFCFSRND	1860
Qy	1861	SLSLDDFDDDDVLSREKALIRKAKENKSEAKVTSHTELTSQOOSANKTAQAKQPINR	1920
Db	1861	SLSLDDFDDDDVLSREKALIRKAKENKSEAKVTSHTELTSQOOSANKTAQAKQPINR	1920
Qy	1921	GQPKPILOKOSTPQSSKDIPIDRGAATDEKLQNFATENTPVCFSHNSLSLSLSDIDQENN	1980
Db	1921	GQPKPILOKOSTPQSSKDIPIDRGAATDEKLQNFATENTPVCFSHNSLSLSLSDIDQENN	1980
Qy	1981	NKENPIKETETPPDSQGEPSKQASGYAPKSFHVHVEDTPVCFSRNNSLSLSISISEDDLQ	2040
Db	1981	NKENPIKETETPPDSQGEPSKQASGYAPKSFHVHVEDTPVCFSRNNSLSLSISISEDDLQ	2040
Qy	2041	ECISSAMFKKKKPRLKCDNEKHSRPNWGGILGDLTLDLKDIQRPDSEHGLSPDSENF	2100

```

Db 2041 ECISAMPKKKPSRLKGDNEKSPRNMGGILGEDTLDLKDQRPDPSEHGLSPDSENF 2100
Qy 2101 WKATOEGANSIVSLHQAARAAACLSRAASDSISILSKSGISLGSFPHITPQOEKPF 2160
Db 2101 WKATOEGANSIVSLHQAARAAACLSRAASDSISILSKSGISLGSFPHITPQOEKPF 2160
Qy 2161 SNKGPRILKPEKSTLTETKIESBKIKGKKVYKSLITGKVRNSEISQWKPLOAN 2220
Db 2161 SNKGPRILKPEKSTLTETKIESBKIKGKKVYKSLITGKVRNSEISQWKPLOAN 2220
Qy 2221 MPSISRGRTWIIHICVNRNSSTSPVSKKGPPLKTPASKSPSGQTATTPRGAKPSVK 2280
Db 2221 MPSISRGRTWIIHICVNRNSSTSPVSKKGPPLKTPASKSPSGQTATTPRGAKPSVK 2280
Qy 2281 ELSVPAQTSOIGSSKAPRSGSRDTPSRPAQPLSRPIQSPGRNSISPGRNGISPPN 2340
Db 2281 ELSVPAQTSOIGSSKAPRSGSRDTPSRPAQPLSRPIQSPGRNSISPGRNGISPPN 2340
Qy 2341 KLSQLPRTSGPSTASTKSSGSKGMYTSPGRQMSQQNLTKQTGLSKNASSIPRSESAS 2400
Db 2341 KLSQLPRTSGPSTASTKSSGSKGMYTSPGRQMSQQNLTKQTGLSKNASSIPRSESAS 2400
Qy 2401 LNMNNGNGANKKVELSRMSTKSSGESDRSRPVLVROSTIKAPSPPLRKLKES 2460
Db 2401 LNMNNGNGANKKVELSRMSTKSSGESDRSRPVLVROSTIKAPSPPLRKLKES 2460
Qy 2461 SFESLSPSRPASPTRSOAQTIVLSPSLPDMSLSTHSSVOAGGWRKLPNLSPTIEYND 2520
Db 2461 SFESLSPSRPASPTRSOAQTIVLSPSLPDMSLSTHSSVOAGGWRKLPNLSPTIEYND 2520
Qy 2521 RPAKHDIASHSPRLPINRSGTWKREHSHKSSSLPRVSTWRTTSGSSSILSASS 2580
Db 2521 RPAKHDIASHSPRLPINRSGTWKREHSHKSSSLPRVSTWRTTSGSSSILSASS 2580
Qy 2581 SEKAKEDEKHVNSISGTSKQENQVSAKGTWKIKENEPSPNTSTQTVSSGATNGAES 2640
Db 2581 SEKAKEDEKHVNSISGTSKQENQVSAKGTWKIKENEPSPNTSTQTVSSGATNGAES 2640
Qy 2641 KTLIYQAPAVKTEDVWVRIEDCPINPRSGSPGTGNTPPVDSVSEKANPNIKDSK 2700
Db 2641 KTLIYQAPAVKTEDVWVRIEDCPINPRSGSPGTGNTPPVDSVSEKANPNIKDSK 2700
Qy 2701 QAKQNVGNGVPMETVGLNRLTSFIQVADPOKGTETIKFQGNPPVPVSETNESPIVERT 2760
Db 2701 QAKQNVGNGVPMETVGLNRLTSFIQVADPOKGTETIKFQGNPPVPVSETNESPIVERT 2760
Qy 2761 PFSSSSSKHSPSGTVAARVTPFNPNPSPRKSSADSTGARPQIPTPVNNNTKRDST 2820
Db 2761 PFSSSSSKHSPSGTVAARVTPFNPNPSPRKSSADSTGARPQIPTPVNNNTKRDST 2820
Qy 2821 DSTESSGTQSPKXHSGLVTSV 2843
Db 2821 DSTESSGTQSPKXHSGLVTSV 2843

```

```

RESULT 8
US-10-267-502-372
: Sequence 372, Application US/10267502
: Publication No US20040071700A1
: GENERAL INFORMATION:
: APPLICANT: Kim, Jaeseob
: APPLICANT: Galant, Ron
: TITLE OF INVENTION: Obesity Linked Genes
: FILE REFERENCE: LSD-07416
: CURRENT APPLICATION NUMBER: US/10/267,502
: CURRENT FILING DATE: 2003-01-27
: NUMBER OF SEQ ID NOS: 439
: SOFTWARE: Patent in version 3.2
: SEQ ID NO 372
: LENGTH: 2845
: TYPE: PRT
: ORGANISM: Mus musculus

```

```

US-10-267-502-372
Query Match .90.2%; Score 13148; DB 12; Length 2845;
Best Local Similarity 90.0%; Pred. No. 0;
Matches 2566; Conservative 113; Mismatches 159; Indels 12; Gaps 10;
Qy 1 MAAASYDQLLAKQVEALKMENSINLROELEDNHNHILTKLETEASNKKVILKQJQSGIIEDEAM 60
Db 1 MAAASYDQLLAKQVEALKMENSINLROELEDNHNHILTKLETEASNKKVILKQJQSGIIEDEAM 60
Qy 61 ASSGQIDLLERLKEINLDSNFPVGVKLRSMKLSYSGRSRGSVSRSGECSPVPMGSPFR 120
Db 61 -TSGQIDLLERLKEINLDSNFPVGVKLRSMKLSYSGRSRGSVSRSGECSPVPMGSPFR 118
Qy 121 RGFVNGSRESTGYLEELKERSILLADLDKEEKDWYAAOLNLTKRIDSPLTENFSL 180
Db 119 RGFVNGSRESTGYLEELKERSILLADLDKEEKDWYAAOLNLTKRIDSPLTENFSL 178
Qy 181 QTDLTRQLEVEARQIRVAMEBQLGTQDMEKRAQRRIARIQOIEKDIRLRLQSQAT 240
Db 179 QTDLTRQLEVEARQIRVAMEBQLGTQDMEKRAQRRIARIQOIEKDIRLRLQSQAA 238
Qy 241 EAERSSQNHKTHGSHDAERQNEGOGVEINMATSGSGSTRMDHETASVLSSSSTHSA 300
Db 239 EAERSSQNHKTHGSHDAERQNEGOGVEINMATSGSGSTRMDHETASVLSSSSTHSA 298
Qy 301 PRRLTSLHGTVMVYISLLSMLGTHDKDDMSRTLLAVSSSQSDSCISMRQSGCLPLLIQLL 360
Db 299 PRRLTSLHGTVMVYISLLSMLGTHDKDDMSRTLLAVSSSQSDSCISMRQSGCLPLLIQLL 358
Qy 361 HGNKXSVLLGNSSGSEARARASAAALHNIIHSPDDKRGREIRVJLHLEQIRAYCETC 420
Db 359 HGNKXSVLLGNSSGSEARARASAAALHNIIHSPDDKRGREIRVJLHLEQIRAYCETC 418
Qy 421 WEWEAEHPGQDQKNPMPAPVEHQICPAVCVLMKLSFDEHRRHMMELGLOQAIALLQ 480
Db 419 WEWEAEHPGQDQKNPMPAPVEHQICPAVCVLMKLSFDEHRRHMMELGLOQAIALLQ 478
Qy 481 VDCEMYGLTNDHYHITLRRYAGMALTNLTGPDVANKATLCMKGCMRALVAQLKSEEDL 540
Db 479 VDCEMYGLTNDHYHITLRRYAGMALTNLTGPDVANKATLCMKGCMRALVAQLKSEEDL 538
Qy 541 QOVTASVLRNLSRADVNSKTLREVGSVKALMBCALEVKESTLKSLSALNLSAHT 600
Db 539 QOVTASVLRNLSRADVNSKTLREVGSVKALMBCALEVKESTLKSLSALNLSAHT 598
Qy 601 ENKADI CAVDGALAFVLTITYSQTNTLAIIESGGGILRNVSLSLIATNEDHRIQIRENN 660
Db 599 ENKADI CAVDGALAFVLTITYSQTNTLAIIESGGGILRNVSLSLIATNEDHRIQIRENN 658
Qy 661 CLQTLQHLKSHSLTIIVSNACGLTNLSARNPKDQALWDMGAVSMKLNLIHSGHMIA 720
Db 659 CLQTLQHLKSHSLTIIVSNACGLTNLSARNPKDQALWDMGAVSMKLNLIHSGHMIA 718
Qy 721 GSAALRNLMANRPARYKDNIMSPGSSLPFLSHVRKOKALEAELDAQHLSSETFDNIDLS 780
Db 719 GSAALRNLMANRPARYKDNIMSPGSSLPFLSHVRKOKALEAELDAQHLSSETFDNIDLS 778
Qy 781 PKASHRSQRHKSQLYGDIYFDNTHDDNRSDFNTGNMTVLSPLYNTTVLPSSSSRGS 840
Db 779 PKASHRSQRHKSQLYGDIYFDNTHDDNRSDFNTGNMTVLSPLYNTTVLPSSSSRGS 838
Qy 841 LSSRSSEKDRSLERERIGLGNYPATENPGTSKRGLOISTTAAQIAKYMEVSAIHTS 900
Db 839 LSSRSSEKDRSLERERIGLGNYPATENPGTSKRGLOISTTAAQIAKYMEVSAIHTS 898
Qy 901 QEDRSSGSTTEHLCVTBERNALRRSSAAHSTHTYNTFTKSENENRPTCSMPYAKLEYKRSS 960
Db 899 QEDRSSGSTTEHLCVTBERNALRRSSAAHSTHTYNTFTKSENENRPTCSMPYAKLEYKRSS 958
Qy 961 NDSLNSVSSNDGYGKRGQMKPSESYESDEDESKFCSYGOYPADLAHKIHSANHMDNDGE 1020
Db 959 NDSLNSVSSNDGYGKRGQMKPSESYESDEDESKFCSYGOYPADLAHKIHSANHMDNDGE 1018

```

QY 1021 LDTPIYSLKYSDQLNGRQSPQNERWAPKHHIIEDEIKQSQRQSRNQSTTYFYVTE 1080  
DB 1019 LDTPIYSLKYSDQLNGRQSPQNERWAPKHHIIEDEIKQSQRQSRNQSTTYFYVTE 1078  
QY 1081 STDDKHLKQFQFHGQECVSPYRGRGANGSTNVRGSHGNQVNSOSLQCEDDYEDDKP 1140  
DB 1079 NTDDKHLKQFQFHGQECVSPYRGRGANGSTNVRGSHGNQVNSOSLQCEDDYEDDKP 1138  
QY 1141 TNYSERVSEEOH-EEBERPNYISKYNEERKHVDQPIDYSLKYATDIPSSQKOSFQFSK 1199  
DB 1139 TNYSERVSEEOH-EEBERPNYISKYNEERKHVDQPIDYSLKYATDIPSSQKOSFQFSK 1198  
QY 1200 SSSGSSSTEHWSSESNTSPSSNAKRONLHPSSAQSRSGOPKAACTCKVSSINQETI 1259  
DB 1199 NSSAQSTKPELSESENAVPPSSNAKRONLHPSSAQSRSGOPKAACTCKVSSINQETI 1257  
QY 1260 QTYCVEDTPICFRCSLSLSSAEDEIGCQNTTQEADSAANTLQIAEIKGIGTRSAEDP 1319  
DB 1258 QTYCVEDTPICFRCSLSLSSAEDEIGCQNTTQEADSAANTLQIAEIKGIGTRSAEDP 1317  
QY 1320 VSEVPANVCHARTSSRLQSSLSSESARH-KAYEFFSGAKSPKSGAQTCKSPPEHYVQ 1378  
DB 1318 ATEVPANVCHARTSSRLQSSLSSESARH-KAYEFFSGAKSPKSGAQTCKSPPEHYVQ 1377  
QY 1379 ETPLMFSRCTSVSSLDSPESRSIASSVQSEPCSGMVSGIISPSDLPDPSPGOTMPPSRSKT 1438  
DB 1378 ETPLMFSRCTSVSSLDSPESRSIASSVQSEPCSGMVSGIISPSDLPDPSPGOTMPPSRSKT 1437  
QY 1439 PPPPQTAQTRKVPKNAKTAERKESGPKQAANVAVQVRQVLPDADTLHPATSTPD 1498  
DB 1438 PPPPQTAQTRKVPKNAKTAERKESGPKQAANVAVQVRQVLPDADTLHPATSTPD 1497  
QY 1499 GPCSSSSLSLSDPESRSIASSVQSEPCSGMVSGIISPSDLPDPSPGOTMPPSRSKT 1558  
DB 1498 GPCSSSSLSLSDPESRSIASSVQSEPCSGMVSGIISPSDLPDPSPGOTMPPSRSKT 1556  
QY 1559 SEKDLSDSDDDDDILEECIISAMPTKSRKKGKPAQATASKLPPPVARKPSQLPVYKLL 1618  
DB 1557 SEKDLSDSDDDDDILEECIISAMPTKSRKKGKPAQATASKLPPPVARKPSQLPVYKLL 1616  
QY 1619 PSQNLQOKHVSFTPGDDMRVYCEGTPIFNSTATSLSDLTIESPNNELAAEGVRGG 1678  
DB 1617 PAQNLQOKHVSFTPGDDMRVYCEGTPIFNSTATSLSDLTIESPNNELAAEGVRGG 1676  
QY 1679 AQSGFEKEDTPTBGRSTDDAQRGKISSIVTPDLDDNKAEEGDILAEICINSAMPKKGSH 1738  
DB 1677 AQSGFEKEDTPTBGRSTDDAQRGKISSIVTPDLDDNKAEEGDILAEICINSAMPKKGSH 1736  
QY 1739 KPFVKKIMDQVQASASSAPKQKQKKTSPVKPIQONTEYTRVRKNAQSKN 1798  
DB 1737 KPFVKKIMDQVQASASSAPKQKQKKTSPVKPIQONTEYTRVRKNAQSKN 1796  
QY 1799 LNAERVFSDNKKONLKNNSKDPNDKLPNNEDRVGRGSAFDSPHHTPIEGTPYCFPSR 1858  
DB 1797 VNTETFDNKKDPSLQTNAKAFNEKLPNNEDRVGRGSAFDSPHHTPIEGTPYCFPSR 1856  
QY 1859 NDSLSLSDPDDDDVLSREKALREKAKNKGSEAKVTSHTELTSNQOQANKTOIAKOPI 1918  
DB 1857 NDSLSLSDPDDDDVLSREKALREKAKNKGSEAKVTSHTELTSNQOQANKTOIAKOPI 1916  
QY 1919 NRGQPKILOKSTPQSKDTPDRGAATDEKLONFALIENTPVCFSHNSLSLSDIDOE 1978  
DB 1917 NRGQPKILOKSTPQSKDTPDRGAATDEKLONFALIENTPVCFSHNSLSLSDIDOE 1976  
QY 1979 -NNKNEPEIKETETPPDSQCEPSKQASGAYAPKSFHVEDTTPVCFSRNSSLSDSDSD 2037  
DB 1977 -NNKNEPEIKETETPPDSQCEPSKQASGAYAPKSFHVEDTTPVCFSRNSSLSDSDSD 2036  
QY 2038 LLOECISSAMPKKKPSRLKGNKHSPRNMGGILGEDTLTDKDIQRPDSEHGLSPDSE 2097  
DB 2037 LLOECISSAMPKKKPSRLKGNKHSPRNMGGILGEDTLTDKDIQRPDSEHGLSPDSE 2096

QY 2098 NFDKAIORGANSTVSSLHQ-AAAAACLSROASSSDSILSLKSGISLGSPHLPDQBE 2156  
DB 2097 NFDKAIORGANSTVSSLHQ-AAAAACLSROASSSDSILSLKSGISLGSPHLPDQBE 2156  
QY 2157 KPFTSNKGRILKPKBSTLETKIESEKIGKGGKVKSLITGKVRNSENISGOMKQ 2216  
DB 2157 KPFTSNKGRILKPKBSTLETKIESEKIGKGGKVKSLITGKVRNSENISGOMKQ 2216  
QY 2217 LQANPESIRGRTHIIPGVNRSSSSTSPVKKGGPLKTPASKSPSEGOTATSPREGAKP 2276  
DB 2217 LQANPESIRGRTHIIPGVNRSSSSTSPVKKGGPLKTPASKSPSEGOTATSPREGAKP 2276  
QY 2277 SVKSELSPVARQTOIGGSSKAPSRSGRSDTSPSPAQOPLSRPIQSPGRNISPGRNGI 2336  
DB 2277 AGKSELSPITRQTSQISGNGKSSRSGRSDTSPSPAQOPLSRPIQSPGRNISPGRNGI 2336  
QY 2337 SPNKLSQLPRSTSPSTASTKSSGSGKMSYTFPGROMSQOONLTQOTGLSKNASSIPRSES 2396  
DB 2337 SPNKLSQLPRSTSPSTASTKSSGSGKMSYTFPGROMSQOONLTQOTGLSKNASSIPRSES 2396  
QY 2397 ASKGLNOMNGNGANKVELSRMSSTKSSGSGSDSERPVLVRQSTFIKEAPSILRRKL 2456  
DB 2397 ASKGLNOMNGNGANKVELSRMSSTKSSGSGSDSERPVLVRQSTFIKEAPSILRRKL 2456  
QY 2457 EESASFESLSPSRPASPTRSOAQTVPVLSPLSPDMSLSTHSSVQAGGWRKLPNLSPTIE 2516  
DB 2457 EESASFESLSPSRPASPTRSOAQTVPVLSPLSPDMSLSTHSSVQAGGWRKLPNLSPTIE 2516  
QY 2517 YNDGRPAKHDIARSHESPSRLPINESCTWKEHSHSSSLPRVSTWERTGSSSSILSA 2576  
DB 2517 YNDGRPAKHDIARSHESPSRLPINESCTWKEHSHSSSLPRVSTWERTGSSSSILSA 2576  
QY 2577 SSESSEKAKSEDEKHVNSISGTQSKENOVSAKGTWRKIKENEFSTNSTQTSVSSGATN 2636  
DB 2577 SSESSEKAKSEDEKHVNSISGTQSKENOVSAKGTWRKIKENEFSTNSTQTSVSSGATN 2636  
QY 2637 GAESKTLIQMAPAVSKTEDVWVRIEDCPINPRSCRSPTGNTPPVIDSVSEKANPNKD 2696  
DB 2637 GAESKTLIQMAPAVSKTEDVWVRIEDCPINPRSCRSPTGNTPPVIDSVSEKANPNKD 2696  
QY 2697 SKDQO--AKONGVNGSVPMRTVGLNRLTSTFIQVDAPQKGTKEIKPGQNNPVPVSEFNE 2753  
DB 2697 SKDQO--AKONGVNGSVPMRTVGLNRLTSTFIQVDAPQKGTKEIKPGQNNPVPVSEFNE 2753  
QY 2754 SPIVERTPSSSSSKHSPSGTVAARVTFNPNPRKSSADSTARSQIPTPVNNT 2813  
DB 2754 SPIVERTPSSSSSKHSPSGTVAARVTFNPNPRKSSADSTARSQIPTPVNNT 2813  
QY 2814 KKDSDKTDSTESSGTSQSPKSHSGSYLVTSV 2843  
DB 2814 KKDSDKTDSTESSGTSQSPKSHSGSYLVTSV 2843  
QY 2816 KKDSDKTDITESSGAQSPKSHSGSYLVTSV 2845  
DB 2816 KKDSDKTDITESSGAQSPKSHSGSYLVTSV 2845

## RESULT 9

US-09-987-482-2  
; Sequence 2, Application US/09987482  
; Publication No. US20020184656A1  
; GENERAL INFORMATION:  
; APPLICANT: BHANDARI, POONAM  
; APPLICANT: SHASHIDHARA, L.S.  
; TITLE OF INVENTION: IN VIVO ASSAY SYSTEM FOR SCREENING AND VALIDATION OF  
; TITLE OF INVENTION: DRUGS AND OTHER SUBSTANCES  
; FILE REFERENCE: 056859-0134  
; CURRENT APPLICATION NUMBER: US/09/987,482  
; CURRENT FILING DATE: 2002-03-21  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 912  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-987-482-2

Query Match 32.4%; Score 4719; DB 9; Length 912;  
Best Local Similarity 99.6%; Pred. No. 8.2e-259;  
Matches 908; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 959 SSNDLSNSVNDGYGRGQKPSIESYSDDESKSCYQYADLAHAKHSANHMDND 1018  
DB 1 SSNDLSNSVSSDGYGRGQKPSIESYSDDESKSCYQYADLAHAKHSANHMDND 60

QY 1019 GELDTPIYSLKXSDQLNSGRSPSQNERWARPKHIIIEDEIKQSEQRNRQSTTPYV 1078  
DB 61 GELDTPIYSLKXSDQLNSGRSPSQNERWARPKHIIIEDEIKQSEQRNRQSTTPYV 120

QY 1079 TESTDDKHLXFOHFQGOECVSPYRSGANGSETNRVGNHGINQNVSLCOEDDYEDD 1138  
DB 121 TESTDDKHLXFOHFQGOECVSPYRSGANGSETNRVGNHGINQNVSLCOEDDYEDD 180

QY 1139 KPTNYSERYSEEEHQBEEERPTNYSIKYNEEKHRVQPIDYSLKYATDIPSSQKQSFSPS 1198  
DB 181 KPTNYSERYSEEEHQBEEERPTNYSIKYNEEKHRVQPIDYSLKYATDIPSSQKQSFSPS 240

QY 1199 KSSGGQSKTEHMSSESSENTSTPSSNAKRONQOLHPSSAQSRSQPOKAATCKVSSINQST 1258  
DB 241 KSSGGQSKTEHMSSESSENTSTPSSNAKRONQOLHPSSAQSRSQPOKAATCKVSSINQST 300

QY 1259 IQYCVEDDTICPSRCSLSLSAEDIEICNQTTOEADSNANTLOIAEIKGIGTRASD 1318  
DB 301 IQYCVEDDTICPSRCSLSLSAEDIEICNQTTOEADSNANTLOIAEIKGIGTRASD 360

QY 1319 PVSEVPVAVSQHPTKSSRLQSSLSSESARHKAVEFPSSGAKSPSKGAOTPKSPPEHYQ 1378  
DB 361 PVSEVPVAVSQHPTKSSRLQSSLSSESARHKAVEFPSSGAKSPSKGAOTPKSPPEHYQ 420

QY 1379 ETPLMFSECTVSASLDSFERSSTASSVQSEPCGMVSGIISPSDLPSQGMTPSPRSKT 1438  
DB 421 ETPLMFSECTVSASLDSFERSSTASSVQSEPCGMVSGIISPSDLPSQGMTPSPRSKT 480

QY 1439 PPPPTQAKTREVPKNAKPTAEKRESGPKQAAVNAQVQVLPDADTLHFATESTPD 1498  
DB 481 PPPPTQAKTREVPKNAKPTAEKRESGPKQAAVNAQVQVLPDADTLHFATESTPD 540

QY 1499 GFSCSSLSALSLEDEPIQKDVLRIMPVQENDNGNETESEQPKESNENQEKAKTID 1558  
DB 541 GFSCSSLSALSLEDEPIQKDVLRIMPVQENDNGNETESEQPKESNENQEKAKTID 600

QY 1559 SEKDLDSDSDDDIEILEECIIISAMPTKSSRKQKPAQATSKLPPPVARKPSQOLPVYKLL 1618  
DB 601 SEKDLDSDSDDDIEILEECIIISAMPTKSSRKQKPAQATSKLPPPVARKPSQOLPVYKLL 660

QY 1619 PSQNRLOPKHVFTPGDDMPRVYCEGTPIINFSTATSLDITSPSPNELAAGEVRGG 1678  
DB 661 PSQNRLOPKHVFTPGDDMPRVYCEGTPIINFSTATSLDITSPSPNELAAGEVRGG 720

QY 1679 AQSGEPFKRTIITEGRSTDEAGCKTSSVTIPELDNKAEEGDIIAECINSAMPKGS 1738  
DB 721 AQSGEPFKRTIITEGRSTDEAGCKTSSVTIPELDNKAEEGDIIAECINSAMPKGS 780

QY 1739 KPRVKKIMQVOQAQASSAPNKNQLDGKKKPTSPVKPIPONTYRTRVRKNADSKN 1798  
DB 781 KPRVKKIMQVOQAQASSAPNKNQLDGKKKPTSPVKPIPONTYRTRVRKNADSKN 840

QY 1799 LNAERVPDNKSKQNLKNNSKDFNDKLPNNEDVRGSAFDSPHHYTPTEGTPCFER 1858  
DB 841 LNAERVPDNKSKQNLKNNSKDFNDKLPNNEDVRGSAFDSPHHYTPTEGTPCFER 900

QY 1859 NDSLSSLDFFDD 1870  
DB 901 NDSLSSLDFFDD 912

RESULT 10

US-09-987-482-3

; Sequence 3, Application US/09987482

; Publication No. US20020184656A1

GENERAL INFORMATION:  
; APPLICANT: BHANDARI, POONAM  
; APPLICANT: SHASHIDHARA, L.S.  
; TITLE OF INVENTION: IN VIVO ASSAY SYSTEM FOR SCREENING AND VALIDATION OF  
; TITLE OF INVENTION: DRUGS AND OTHER SUBSTANCES  
; FILE REFERENCE: 056859-0134  
; CURRENT APPLICATION NUMBER: US/09/987,482  
; CURRENT FILING DATE: 2002-03-21  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 767  
; TYPE: PRF  
; ORGANISM: Homo sapiens  
US-09-987-482-3

Query Match 26.7%; Score 3889; DB 9; Length 767;  
Best Local Similarity 99.9%; Pred. No. 7.7e-212;  
Matches 766; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAASYSQDQLLKQVEALKXWNSNLRQLEEDNSNHLTKLETEASNKKEVLKOLQGSIEDEAM 60  
DB 1 MAASYSQDQLLKQVEALKXWNSNLRQLEEDNSNHLTKLETEASNKKEVLKOLQGSIEDEAM 60

QY 61 ASSGQIDLLERLKEINLDSSNFPQVKLRKWSLSYSGSRGVSRSRSGECSVPVMSGFPR 120  
DB 61 ASSGQIDLLERLKEINLDSSNFPQVKLRKWSLSYSGSRGVSRSRSGECSVPVMSGFPR 120

QY 121 RGFVNGSRESTGYLEELKERSLLADLDKEEKQWYQALQNLTKRIDSLPTENFSL 180  
DB 121 RGFVNGSRESTGYLEELKERSLLADLDKEEKQWYQALQNLTKRIDSLPTENFSL 180

QY 181 QTDLTRQLEVEARQIVAMEEQGTQDMKRAQRIARIQIQEKDILRIQLQSQAT 240  
DB 181 QTDLTRQLEVEARQIVAMEEQGTQDMKRAQRIARIQIQEKDILRIQLQSQAT 240

QY 241 EAERSSQNKHETGSHDAERQNEGQVGEINNATSGNCGSTTRMDHETASVLSSTHSA 300  
DB 241 EAERSSQNKHETGSHDAERQNEGQVGEINNATSGNCGSTTRMDHETASVLSSTHSA 300

QY 301 PRRLTSHLGTKEVWYVLSLMLGTHDKDDMSRTLLANSSQSDSCISMRQSGCLPLLIQLL 360  
DB 301 PRRLTSHLGTKEVWYVLSLMLGTHDKDDMSRTLLANSSQSDSCISMRQSGCLPLLIQLL 360

QY 361 HGNDKDSVLLGNSRSGSKEARASAAALHNIHSOPDDKRGREIRVLHLLLEQIRAYCETC 420  
DB 361 HGNDKDSVLLGNSRSGSKEARASAAALHNIHSOPDDKRGREIRVLHLLLEQIRAYCETC 420

QY 421 WEMOEAHEPGMDQKMPAPVEHQICPAVCVLMKLSFDEEHRHAMNELGGLQAIALLQ 480  
DB 421 WEMOEAHEPGMDQKMPAPVEHQICPAVCVLMKLSFDEEHRHAMNELGGLQAIALLQ 480

QY 481 VDCSMYGLTNDHYSITLRRYAGVATLNLTFQDVANKATLCSMKGCMALVAQLKSEEDL 540  
DB 481 VDCSMYGLTNDHYSITLRRYAGVATLNLTFQDVANKATLCSMKGCMALVAQLKSEEDL 540

QY 541 QQVIASVLRNLSWPAADVNSKKTUREVGSVKALMECALEVKESTLKSVALNLSAHCT 600  
DB 541 QQVIASVLRNLSWPAADVNSKKTUREVGSVKALMECALEVKESTLKSVALNLSAHCT 600

QY 601 ENKADIICAVGALAFVGLTLYRSQNTLAIIESGGGILNRVSSLTATNDEHQIILRENN 660  
DB 601 ENKADIICAVGALAFVGLTLYRSQNTLAIIESGGGILNRVSSLTATNDEHQIILRENN 660

QY 661 CLQTLQHLKSHSLTIYSNACGTLNLSARNPKDQEAALWDMGAVSMKNLIHSHKRMIAM 720  
DB 661 CLQTLQHLKSHSLTIYSNACGTLNLSARNPKDQEAALWDMGAVSMKNLIHSHKRMIAM 720

QY 721 GSAALRNANRPARYKDNIMSPGSSLSLHVRRKQKALEAELDAQ 767  
DB 721 GSAALRNANRPARYKDNIMSPGSSLSLHVRRKQKALEAELDAQ 767

```

RESULT 11
US-10-267-502-373
; Sequence 373, Application US/10267502
; Publication No. US20040071700A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jaeseob
; APPLICANT: Galant, Ron
; TITLE OF INVENTION: Obesity Linked Genes
; FILE REFERENCE: LSD-07416
; CURRENT APPLICATION NUMBER: US/10/267,502
; CURRENT FILING DATE: 2003-01-27
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 373
; LENGTH: 2274
; TYPE: PRI
; ORGANISM: Mus musculus
US-10-267-502-373

Query Match      24.1%; Score 3508.5; DB 12; Length 2274;
Best Local Similarity 34.4%; Pred No. 1.4e-189;
Matches 1005; Conservative 356; Mismatches 789; Indels 773; Gaps 98

QY 4 ASYDQLLKQVEALKWKNENLNLSQELNEDSNHLTKLETFASNKWEVLKQLOGSIEDFA--MA 61
Db 6 ASVEQLVRQVEALKAEKNTLHLSQELRDNSHLKLETETSGMKEVLKHLQKLEQEARVLV 65
QY 62 SSGQIDLLERLKELNLDSSNFPGVKLRSKXLSLASYSGREGSVSSRGECSPYPMGSPFPR 121
Db 66 SSGQTEVLEQLKALQTDLSYYLNKLFHAP----ALGPEP---AARTPEGSPV-HGSGPSK 117
QY 122 -GFVNGSRESTGYLZELEKERSLLIALDKEEKEKWYYAQLQMLTKRIDSPLTENFSL 180
Db 118 DSFGELSRATIRLLEELDQERCFLLSBEKEKEKLWYYSQLOGLSKRLDELPHVDFTSM 177
QY 181 QTDLTRQLREYARQIRVAMEEQIGTQDNKGAQRRIARIQIEKDIL-----RIQQ--- 233
Db 178 QMDLIRQLLEPAQRIHSLMEERFCTSDENVQRAIPASLEQIDKELLEAQDRVQOTEP 237
QY 234 ---LLOSQATEAERSQNHETGSHDAERQNEGQVGVEINMATSGNGQSTTRMDHETASV 291
Db 238 QALLAVKPVAVEEQEABVPHEDGTPQ-----PGN----- 269
QY 292 LSSSTHSAPRLTSHLGTQKVMYVYLLSMLGTHDKDDMERTLLAMSSQDSCISNROSG 351
Db 270 -----SKVEVVFLLSMLATRDQEDTARTLLAMSSSPESCVMARRSG 311
QY 352 CLPLLIQLLHGNKDQSV---LNLGNSRGSKBARAPASAALENITHSDPPDKGRRIEVLH 408
Db 312 CLPLQLQLHGTGFEAGSVGRAGIPGAPGAKAKARMANALNITHVSPDQGLARKEMRVLH 371
QY 409 LLEQIRAYCETCWEOQEAHEFGMDQNPAPVVEHQICPAVCVLMKLSFDEEHRHAMNE 468
Db 372 VLEQIRAYCETCWDMLQARDSGTE----TPVPIEPQICQATCAVMKLSFDEEYRRAMNE 426
QY 469 LGGLOAIAELLQVDCMYGLTNDHYSITLRYAGMALTNITFGDVANKATILCSMKGCWRA 528
Db 427 LGGLOAIAELLQVDEYEMHKMTRDPLNALRLARYAGMTLNLITFGDVANKATILCARRGMEA 486
QY 529 LVAQLKSSEDLQQVIAVLNLSWRADVNSKKTLEVGSVKALMECALEVKESTLKSIV 588
Db 487 IVAQLGSSEELHQVVSILNLSWRADINSKKVLRVSGNTALMECVLRASKESTLKSIV 546
QY 589 LSAIWNLSAHTENKADICAVDGALAFVLGTLTYRSQNTNLAIIESGGILNRNVSLSIAT 648
Db 547 LSAIWNLSAHTENKAAICQVDGALGALVSTLTLYRCQGNSLAVIESGGILNRNVSLSIAT 606
QY 649 NEDHRQILRENNCLQTLLQHLKSHSLTIVSNACCTIWNLSARPKDQEAALWDMGAVSLK 708
Db 607 REDIYQVLRDHNCLQTLLQHLTSHSLTIVSNACCTIWNLSARPKDQEAALWDMGAVSLK 666
QY 709 NLIHSKHQVIAMGSAALNLRNMANRPKAYKDANI-KSPGSSLESPLHVRKQKALEAELDAQ 767

```

```

QY 1813 KNLKNNKDFNDKLNNEEDVRGSPAFDPSPHVHTPIEGTFYCFPSRNDLSLSLDFDDDDV 1872
Db 1467 CQT-----RTGDGALQSLCTTTEZAVYCF-----YSDDE 1499
QY 1873 DLS-----REKALRKAENKESAKVTSHTLTNSQOSANKTOIAKQPIN---RGOP 1923
Db 1500 PEATAPPRASAIPLAKKEKAGK-----ETPGRAAQATPLPVRAP 1544
QY 1924 KPILQKOSTPQSSKDIPDRGAATDEKLNFAIENTVPCFSHNSLSLSDIDQENNKE 1983
Db 1545 RLI-----VDETPPCYSLTSSASSLSE----- 1566
QY 1984 NEPIKETEPPDSQGEPSKQASGYAPKSFHVEDTPVCFERNSSLSLSDSEDDLQECI 2043
Db 1567 -----PEA---PEQPANHARGPEQSGKQD-----SGPSRAEBELLQRCI 1603
QY 2044 SSAMP-----KKKPSRLKGDNEKSPRNMGILGEDLTLDLKDIOQPDSE-HGLS 2093
Db 1604 SLAMPRTQTVPGRRRKPRALRS-----IRPTEIT-----OKQOEYVAGSD 1646
QY 2094 P--DSENFDMKAIQEGANSIVSSLIHQAAAACLSROASSDSLSILKSGISLGSPPHLT 2151
Db 1647 PASDLSVEWQAQEGANSIVTWLHQAAKASL--EASSEDLSLSIVSGVSGSTLQ-- 1702
QY 2152 PDQEKPTNKGPRILKPKGEKSTLETYKI--ESESIGIKGGKKVYKSLITGKVR--SNSEI 2209
Db 1703 -----PSKLKRGKPAEAGGAWRPEKRGTTSTK-----INGSPRLNGPE 1743
QY 2210 SGQMKQPLQANPISIGRTMIHIPGVNSSSTSPVSKGGLPKTPASKPSPEGOTATT 2269
Db 1744 KAKGTQMGAGESTMLRGRTIVY-----SAGPASRTQSG--ISGPTTPKTGTSGTT 1795
QY 2270 SPRGAKPSVXSELSVPVARTQSOIGGSKAPSGSRDSTPSRPAQOPLGRIPQSPORNSI 2329
Db 1796 QPETV-----TKAPSPQQRSLRHPGKISLAALRHPPRSAT 1834
QY 2330 SPGRNGISPNKLSQLPRTSSP-----STASTKSG-----SGKWSYTSQROMSQNLTK 2380
Db 1835 PPAKATPSSSSSQTSPASQPLPRRPLATPGGLPGPGGLVPKSPARAL-----LAK 1890
QY 2381 QTGLSKNASSI-----PRSESASKGLNQMNNGA--NKKVELSRMSSTK 2423
Db 1891 QHKTQKSPVRIPFMORPARRVPPPLARPSPGSRGAGAEPTGARGSLGLVRMASAR 1950
QY 2424 SSGSESRSRPLVLRSTIKAPSTLRRKLEESAPESLSPSRPASPTRSOATPV 2483
Db 1951 SSGSES--SDRSGFRLLTIIKESPG--LLRRRSSELSADSTASTQASPRGRPALPA 2007
QY 2484 L--SPSLPMWLSLTHSSVOAGGRKLPPLNLSPTIEYDNGRPAKRHDIAARSHSESRLP 2540
Db 2008 VFLCSSRCDLRSVPROPLAQRSPQAKPGLAPL-----APR-----RTSSESRLP 2055
QY 2541 INRGTWKRSHKSSSLPVRSTWRTTGGSSSLSASSSESEKAKEDKHVNSIS--GT 2598
Db 2056 V-EASPGRPETVKRYASLPHISVRSRSDSAVSPTTOANATRRGSDGEARPLPRVAPPT 2114
QY 2599 KQSKENQVSAGTWRKIKENEFS-----PTNSTSTQVSSGATNGAESKTLIYQMAPV 2651
Db 2115 -----TWRRIKDEVDPHILRSTLTPALPLRVSSPEDSPAGTPO----- 2153
QY 2652 SKTEDVVRIEDCPINNPRSGRPT--GNTPPVIDSVEKANPNIKSKDNQAKQNVNG 2709
Db 2154 RKTSDAVVQTEDVATSKTNSSTSPLESRDPP-----QAP-----A 2189
QY 2710 SVPMRTVGLNRLTSFTQVADPQKGTIEIKPGQNNPVVSETNESPTVERTPSSSSSK 2769
Db 2190 SGPVAPQSGD-----VDGP-----VLTKPPASAPFP--HEGLSAVIAGFP-----TSR 2230
QY 2770 HSPSGTVAARVTPFNPNPSRKS--ADSTSRAPQIPTPVN 2810
Db 2231 HGSPSR--AARVPPFNVPFMAAATWASDSAVEKAPVSSPAS 2271

```

```

RESULT 12
US-10-267-502-371
; Sequence 371, Application US/10267502
; GENERAL INFORMATION:
; APPLICANT: Kim, Jaeseob
; APPLICANT: Galant, Ron
; TITLE OF INVENTION: Obesity Linked Genes
; FILE REFERENCE: LSD-07416
; CURRENT APPLICATION NUMBER: US/10/267,502
; CURRENT FILING DATE: 2003-01-27
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 371
; LENGTH: 2303
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-267-502-371

Query Match      23.4%; Score 3414.5; DB 12; Length 2303;
Best Local Similarity 34.3%; Pred. No. 3.1e-184;
Matches 997; Conservative 374; Mismatches 810; Indels 723; Gaps 97;

QY 2 AAASVDQLLKOVEALKWENSRLROLEDNSNHLTKLEASNMKEVLKQOGSTEDEA-- 59
Db 4 SVAPYEQLVRQVEALKKAENSHLROELNDNSHLSKLETSIGMKEVLKHLQGLQEARV 63
QY 60 MASSQCIDLLERLXELNLDSSNFPVKLRKMSLRYSYSGREGSVSSRGECSPVPMGSP 119
Db 64 LVSSQTEVLQKALQMDITSLYNLKFQPP---TLGPEP---AARTPEGSPV-HGSGP 115
QY 120 RR-GFVNGSRSTGYLBELEKERSLLADLDKEKEKOWYTAQONLTUKLIDSLPLTE-N 177
Db 116 SKDSFGELSRATILBELDRERCFLNEKEKEKMLWYQSQGLSKLDELPHVETQ 175
QY 178 FSLQDTLRRQLEYEAQIRVAMEEQELGTQDMKRAQRRIARIQIEKDIRIROLLOS 237
Db 176 FSMQWDLIRQLEFEAQHIRSMEERFGTSDMVQRAQIRASRLQEQIDKELLE----- 228
QY 238 QATSAERSSQNHETGSHDAERQEGQGVGEINWATSGNGQSGSTRMDHETASVLSSST 297
Db 229 -----AQDRVQQTPEQALLAV-----KSPVDEDPETEVPTHPE 262
QY 298 HSAPRLTLHLGTKEVYVYSLMLGTHDKDDMSRTLLAMSSSDSCISMRQSCLEPLLI 357
Db 263 DGTPOGNS-----KVEVFWLLSLMTRDQEDTARTILAMSSSPESCVAVRRSGCLPLL 318
QY 358 QLLHGNKDS--VLLGNRSGSKAPAPASAAALNIIHSQPDCKRGRREIRVHLHLQIR 414
Db 319 QILHGTAAAGRAGAPGAPGAKDARMANAALNIVFSQPDQGLARKEMRVLHVLEQIR 378
QY 415 AYCECTEMQEAHPGVD-QDKNPMPAPVEHQICPAVCVLMKLSFDEHEHANNELGLO 473
Db 379 AYCECTDWLQARQGGPEGGGAGSAPIPIEPQICQATCAVNMKLSFDEYRAMNELGLO 438
QY 474 ATAEILLQVDCMYGLTNDHYSTILRRYAGKALTNLTFTGDVANKATLCSMRGCMRALVAOL 533
Db 439 AVAELLQVDYEMHMTDRDPLNLALRRVAGMTLTNLTFTGDVANKATLCARRGCMBAI 498
QY 534 KSEEDLQOVTASVLRLNLSWEADVNSKKTIREVSKVLMCALEKVEKSTLKSLSALW 593
Db 499 ASDSEELHQVYSSILRLNLSWEADINSKVKVUREAGSVTALVQCVLRATKESLKSLSALW 558
QY 594 NLSAHTENKADIICAVDICALFLVLTITYSQTNLTALIESGGGILNRVNSLTIATNEDHR 653
Db 559 NLSAHTENKAAICQVDGALGFLVLTITYKQSNLAIIESGGGILNRVNSLTIATNEDHR 618
QY 654 QILRENNCLQTLLOHLKSHSLTIVSNACGLTWNLSARNPKDOEALWDNGAVSMKLNLSHS 713
Db 619 QVLRDHNCLQTLLOHLTSHSLTIVSNACGLTWNLSARNPKDOEALWDNGAVSMKLNLSHS 678
QY 714 KHKMTAMGSAALRNLMANRPAYK-DANIMS PGSSLPSSLRVHKQKALEAELDAQHLSET 772

```





RESULT 13  
 US-10-267-502-368  
 ; Sequence 368, Application US/10267502  
 ; Publication No. US20040071700A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kim, Jaeseob  
 ; APPLICANT: Galant, Ron  
 ; TITLE OF INVENTION: Obesity Linked Genes  
 ; FILE REFERENCE: LSD-07416  
 ; CURRENT APPLICATION NUMBER: US/10/267,502  
 ; CURRENT FILING DATE: 2003-01-27  
 ; NUMBER OF SEQ ID NOS: 439  
 ; SOFTWARE: Patent in version 3.2  
 ; SEQ ID NO 368  
 ; LENGTH: 2417  
 ; TYPE: PRT  
 ; ORGANISM: Drosophila melanogaster  
 US-10-267-502-368

Query Match 12.13; Score 1763; DB 12; Length 2417;  
 Best Local Similarity 25.34; Pred. No. 1.6e-90;  
 Matches 720; Conservative 326; Mismatches 836; Indels 964; Gaps 101;

QY	243	ERSSONKHETGSHDAERON-EGQGVGEINMATSGGQSTTR-----MDHETASVLSSS	295
DB	129	ELREWRHSILDRNFERQSQOQQQLDELPPRGGGSPASAGRPSRKEPSYTLRSFLDGD	188
QY	296	STHSAPR-----RLTSH-----LGTKVMYVYLLSMLGTHDKDMSRTLLAM	337
DB	189	APAPAPRLPKGAAMTTFGDERYTSVAEATLGSKEVCYVILLSMLGNSNDPLEMAKFLLEL	248
QY	338	SSSQDCISWRQSCPLPILLQLHGNKDSVLLNGSEKSEARASAAALHNIHSQPD	397
DB	249	SGNAQSCATLRRSCMLPVQMHPAND-----QEVKRCAGQALHNVVHSHPOE	298
QY	398	KRGREIRVLHLLQIRAYCETCWEQEAHEPGM--DQDNMPAPVVEHQCIPAVCLMK	455
DB	299	KAGREAKVRLDQIVDYCSFLKTLQSGEALADSDRPL-----AAISSLMK	349
QY	456	LSPDEERHAWNELGGLQATABELLOVDCENYV-LTNDHYSITLRRYAGHALTNLTFGDVA	514
DB	350	VSPDEERHAWCELGAHALPNIHLVHDVHVGPKPEDQCCNSLRRYALMALTNLTFGDEN	409
QY	515	NKATLCMKGMRALVAQLKSESDLOQVIASVLRNLSWRADVNSKTLREVGSVKALME	574
DB	410	NKALLCQKQFMALVAQLDSAPDDLLOVTSVLRNLSWRADSNMKAVLNEIGTVTALAL	469
QY	575	CALEVKESTLKVLSALWNLASHCTENKADI CAVDGAALFVGLTYTSRQNTLALIES	634
DB	470	AAMNRSNTLKAILLSALWNLASHCTENKADFCVAVDGAALFVGLMSEYSPSKTLXIEN	529
QY	635	GGGILRVSSLIATNEDPHROILRENNGLOTLLOHLKSHSLTIIVSNACGTLWNLARNPKD	694
DB	530	AGGILRVSSHIAVCEPYRQLLQHNCALILLQOLKSESLTVVSNCGTLWNLARSASD	589
QY	695	QEALWDNGAVMLKNLHSHKXMIAMGSAALRNLMANRPAYK-----DANIMSPG-SSL	749
DB	590	QKFLWDNGAVMLRSLHSHKXMIAMGSAALRNLMANRPAYK-----DANIMSPG-SSL	649
QY	750	PSLHVRKQKALELDAQHLSETFDNLNLSPKASHRSKORHKQSLGYGVYVDTNRHDDN	809
DB	650	PTLEAKAKALQELGEBRHTAETCDNLD-----KLDKERASSSSRRHP-----	677
QY	810	RSDNFTGNMTVLSPLYNTLTVLPSSSSSRGSLDSRSEKDRSLERERGLGNYHPATEN	869
DB	678	-----TGG-----KLDKERASSSSRRHP-----	695
QY	870	PTSSKRGLOISTAGIAKVMEEVSAIHTSQDRSGSGSTTELHCYVTDENALRRSSAAH	929
DB	696	-----APLPLR-----SAMLTKSSSRDSVYSKASDCAVD--HLIRGASASD	734

QY	930	THSNYNTFTKSENSNRCTCMPYAKLEYKRSNDSLSNVSSNDGYGKRGQMKPSIESYSD	989
DB	735	AHR-----KVKPKITDFDLE	749
QY	990	DESKFCSYQYPADLAHLKHSANEMDNDGSLDTPINYSKYSDQLNSGQSPSQNERW	1049
DB	750	ME-----QUTEATEQPIDYSVKIS-----	769
QY	1050	ARPKHIIIEIKQSEORQSRNQSTTPYVYTESTDDKHLKFQPHFGQCEVCVPSYRSRGANG	1109
DB	770	-----ENAKTSTY-----	778
QY	1110	SETNRVSGNHGINQNVOSLQEDDYEDDKPTNYSEYSEEQHEEERPTNYSIKNEE	1169
DB	779	-----QETDL--DQPTDFSLYAE-----NOTESDLDISGPAGGQ	811
QY	1170	KRHVDQPIDYSLKYATDIP--SSQKQSFSSKSSSQSGSKTEH-----MSSSESTSTPSS	1223
DB	812	KSTITPP-----AETVPEKSEGOEILLDDSVKCYQTEDTPYVISNAASVTLRVAA	864
QY	1224	NKRONQLHPS--SAQSRSGOPO--KAATCKVSSINOETIOTYCVEDTPICFRCRCSLS	1278
DB	865	KADAEAYKPEVREVTSKGAPKPLKLSQCGSGSYTPKEPINYCEGTGPGYFGRYDLS	924
QY	1279	SLSSAEDSIG--CNOTTQEADSANTLQIAEIKGIGTSAEDPVSSEVPAVSQHPRTKSSRL	1337
DB	925	SL-----DESGKANQITVGTD-----ADIKPLEKEQE-----	954
QY	1338	QGSLSSSASARKHAFVPPSGAKSPKGAQTPKSPPHYVOETPLMFSRCTSVSSLSDFE	1397
DB	955	-----SQPAEQVLTKPPTQANS-----ALETPLMFSRSSMDSLVHDP	992
QY	1398	SBSIA-----SSVQSEPCSGMVSGIISPSDLPSPGOTMPSRSKTPPPPTQATKREV	1452
DB	993	DVDVANCDDKSSVVD--FSRLASGVISPEIPDPTQSMPOS-----	1033
QY	1453	PKNKAPTAEKESGPKQAAVNAVQVLPDADTLHFFATESTPDGFCSSLSALS--	1511
DB	1034	PRNSVAGSQNVDSPPVWIPASLOPLRSVFE--DLSFNVNHTPAQFSTATSLNSIV	1092
QY	1512	DEPIQKXVLRIMPVQENDNGNE-----TESEQPKESNQEKE--AKT	1556
DB	1093	DDE-----KAPASVAEEDNEDELLANCIINMGKRPTEAVKSTVNVSEVDVAET	1143
QY	1557	IDS-----EKD-----LDD--SDDDI--	1572
DB	1144	IRSYCTEDTPALLSKVPSNTNLSVISMSSTDPKDATAGQAQMAHQSLDSDVSSNASDCG	1203
QY	1573	---EILECIIISAMPTKSRKKGKPAQATASKLPPPVARK--PSQLPVYKLLPSQNRLOPQK	1628
DB	1204	ASGHLLQOCIRDM-----KKPLGEATSDPTAMLRGNGNELPGY--LPS-----	1245
QY	1629	HVSTPGDMPRVYCVGCTPINFSTATSLDLTIES-----PPNELAAGEVGRCQAQSGEF	1684
DB	1246	-----ADEKXK--FLVEDSPCNFVSVGLNLTGVSSLVGPAVLKETE--PSSADQNPEN	1297
QY	1685	EKRPTIPEGSTDEAQGGKTSVTIPELDDNKABEGDILAE-----	1727
DB	1298	KAKPKQEQVRRPFWHQDDSLSSLSIDSEDDTNLLSQAIAGCNRPKSNLGFSSNGKRSS	1357
QY	1728	-----INSAMEKGSKHKPFVKKIMQVQOASASSAPKNQO-----LDG--	1767
DB	1358	SLSSSQPIAANAATSSASLNSAMTVRKSSQOQESVSSVSDSDNDNQSKSIFELCILKGY	1417
QY	1768	KKKKP-----TSPVKPIQNTYEYF-----RVRKXADSKNLNA	1801
DB	1418	KTKEPGARAQQMQBPVIGSSVQSNPSLKQFSLPVQLPSSGGQVKQRHHHHHRRER	1477
QY	1802	ERVPSDNKDSK--KQNLKNN--SKDFNDKLPNNEDEVGSAFSDSPHYTPIETPYCFSR	1858
DB	1478	ER-----ERRDEKLLQECINTGISKKIN--AVPKNV--LATSAAALEPCH--PMAAT--	1523
QY	1859	NDLSLSLDFDDDDVDLGRKAEALFKAKENKESAKVTSTHTSLTSNQOSANKTOAIKQPI	1918



Db 1524 TSASALSTAPDV-----EOKAHATSNPQ---QSSHPSSHILNPI 1563  
Qy 1919 -----NRGQPKPILOKQSTPQSQKOIPDRGAATDEKLNPALEN 1558  
Db 1564 DAIATVTDIVRSPAAPNGNGNSQNGLET-ATGSKOLDSDRSDSNQSFIMETVRL 1622  
Qy 1959 -----TPVCF-----SHNSLSLSLSDIDQENN 1980  
Db 1623 DSALNETICISASEKHDPDLMLKSVBLTMEFTVSAEQLRSSSHNSS-----NSHK 1677  
Qy 1981 KNEPIKETEPSPQSQGPKQAGVAPKSFHVEDTPVCFNRSSLSLSDID---SEDD 2037  
Db 1678 NSSNNTWNESTCPNDVSPFSVQTA-----PV-----LASLSLDEDATEAR 1718  
Qy 2038 LLOECISSAMPKKKPSRLKGDNEK-----HSPRNGGILGDLTLDLKDQIPDSEHGLS 2093  
Db 1719 SLHELIELTPNEQOPESEGETDLVNGHADSVSGGGGLNFQJ-----GGQVQAGVRL 1775  
Qy 2094 PDSENFWDKAIQEGANSIVSSJHQAACLSRQASDSD-----SLSL-----KS 2140  
Db 1776 FORLLFNCTASIMNTNMTIAFEAPALAEENLLQPAATDDDTTMTFSLNSLDLONIRPPS 1835  
Qy 2141 GI-SLGSFP--HLTPDQEKPTSNKGRILPKGEKSTLETKTIESESKGKGGKVKYS 2197  
Db 1836 GMESLNSCYODHSQSSSLRQAMPS--KSPRFARKMPPANLVARALGHUAG----- 1884  
Qy 2198 LITGVRNSETISGOMKOP-----LOANMPTISGRMTIHIPGVNRSSSTSPVSKKGP 2252  
Db 1885 -SAESVNSCNLLDNKPPSLMDELDDSMISVDSIQSEVADGEQDCSMATTISVNYETA 1943  
Qy 2253 L-----KTPASKPSBEGCTATTTPRGAKPSVKSEL----- 2282  
Db 1944 ACDDQMTVLOSCEDEDEATMNDYSAAESTPKHGSTPSNRRSLTPKDKERLTKDRFKT 2003  
Qy 2283 -----SPVARTSQI-----GGSKAPSRSGSRDSTPSRPAQQLSRPI----- 2321  
Db 2004 YTIATSCWEAPEANETLOIEVEAAVPVATPSPRANGRRGSAERYKTQILIECLALIQ 2063  
Qy 2322 -----KTPASKPSBEGCTATTTPRGAKPSVKSEL-----QSP--GRNSISPR 2333  
Db 2064 PQDDCPSEQLSSIRAMQOFTFITDINIGHSQETCESTDHPDAGESPECQONSETSC 2123  
Qy 2334 NG-----ISPPNKLSQL-----PTSSPSTASTKSSGSGKMSYTPSGRQMSQOQLTKQT 2382  
Db 2124 DQOEPDQPPPPSIVDLRTSVVKFTLEPATAVKLVRGKKPAYVSPYMSQSRN----- 2178  
Qy 2383 GLSKNASSIPRESASAKGLNQMNNGANKKVELSRMSTKS-----SGSESDSRSPV- 2436  
Db 2179 -----SNNAPSKKTLPTIAKRSVLPGSGVRLPAKKKTP 2216  
Qy 2437 -----LVROSTIKAPSTLRRKLEASAFESLSPSRPASPTRSQAQTPVLSPSL 2488  
Db 2217 PPEPARLERQCTFKD-----EPTNSNVQVPVVE-TK 2249  
Qy 2489 PMSLSTHSSVQAGWRKLPNLSPTIEYNDGRPAKRDHIAHSSESPSRLPINRSGTWK 2548  
Db 2250 PAQTSPTHRA-----SKLPTKKGTA---SGSPSK-----ACSPKRIPLAPA---R 2289  
Qy 2549 REHSKHSSSLPRVSTWRRTGSSSLSASSESEKAKSEDEKVNISGTSKQENQVSA 2608  
Db 2290 RMTQORANTSLAAGKSHAASRVSGSVSTTPPSRNSNLNGSAAAAAAKINHAQS 2349  
Qy 2609 K-GTWKRIKNEFPSTNS---TSQTVSSGATNGAESKTLIYQMAPAVSKTEDVWRIED 2663  
Db 2350 RIANIWKRVDEAKTKOSSNLNLTQTKTSSNMLNANGTKFTLR-----SSTFD----- 2397  
Qy 2664 CPINNPRSGRSPGTGTPPVIDSVSEK 2689  
Db 2398 -----NTPSTAGGVKSK 2409

RESULT 14

US-09-915-307-5  
; Sequence 5, Application US/09915307  
; Publication No. US20020015943A1  
; GENERAL INFORMATION:  
; APPLICANT: Bienz, Marian  
; TITLE OF INVENTION: Assays, methods and means relating to the modulation of  
; TITLE OF INVENTION: levels of nuclear beta-catenin  
; FILE REFERENCE: SMW/FP5864806  
; CURRENT APPLICATION NUMBER: US/09/915,307  
; CURRENT FILING DATE: 2002-11-22  
; PRIOR APPLICATION NUMBER: US 60/221,892  
; PRIOR FILING DATE: 2000-07-31  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 325  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-915-307-5  
  
Query Match 11.3%; Score 1640; DB 12; Length 325;  
Best Local Similarity 99.4%; Pred. No. 8.3e-85;  
Matches 323; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
Qy 1263 CVEDTPTCSRCSLSLSLSAEDIEICNQTOTQADSANTLQIAEIKGIGTRSAEDPVSE 1322  
Db 1 CVEDTPTCSRCSLSLSLSAEDIEICNQTOTQADSANTLQIAEIKGIGTRSAEDPVSE 60  
Qy 1323 VPAVSQHPRTKSRLOGSSLSSESARHKAVEFPFGAKSPSKGAQTPKSPPEHYVQETPL 1382  
Db 61 VPAVSQHPRTKSRLOGSSLSSESARHKAVEFPFGAKSPSKGAQTPKSPPEHYVQETPL 120  
Qy 1383 MFRCTSVSLDSFESRSTASSVQSPFCGMSVGIISPSDLPSGQTMPPPSRKTTPPPP 1442  
Db 121 MFRCTSVSLDSFESRSTASSVQSPFCGMSVGIISPSDLPSGQTMPPPSRKTTPPPP 180  
Qy 1443 PQAQTKREVPKNAKTAERKREGPKQAANAAVQVQLPDPADTLLHFATESTPDGFSC 1502  
Db 181 PQAQTKREVPKNAKTAERKREGPKQAANAAVQVQLPDPADTLLHFATESTPDGFSC 240  
Qy 1503 SSSLSALSLEPTQKDVLRIMPPVQENDNGNETSEQPKESNENQKEAKETIDSEKD 1562  
Db 241 SSSLSALSLEPTQKDVLRIMPPVQENDNGNETSEQPKESNENQKEAKETIDSEKD 300  
Qy 1563 LLDSDDDDDIEILEECIIISAMPTKS 1587  
Db 301 LLDSDDDDDIEILEECIIISAMPTKS 325  
  
RESULT 15  
US-10-029-386-33090  
; Sequence 33090, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G  
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO  
; FILE REFERENCE: ACOMICA-X-2  
; CURRENT APPLICATION NUMBER: US/10/029,386  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 34288  
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 31090  
; LENGTH: 1633  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AC027307.3  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.85  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.84  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.8

OTHER INFORMATION: SWISSPROT HIT: Q61315, EVALUE 1.00e-27  
US-10-029-386-33090

Query Match 10.1%; Score 1474; DB 14; Length 1633;  
Best Local Similarity 26.8%; Pred. No. 2.3e-74;  
Matches 591; Conservative 263; Mismatches 658; Indels 690; Gaps 82;

706 MLKNIHSHKHVIAWGSAALRNLMANRPARYK-DANIMSPGSLPSLHVHVKQKALEAEL 764  
DB 1 MLRNLVHSHKHVIAWGSAALRNLMANRPARYK-DANIMSPGSLPSLHVHVKQKALEAEL 60  
765 DAHLSETFDNIDNLSKASHRSKQ-----RHQSLYGYDYVDTNHDNRSDN-----813  
DB 61 DARHLAQLAHLERKQGPAAEAATKPLPLRHLDGLAODYASDGSDFDDDDAPSSLA 120  
766 FNTG-----NMVLSPYLNTVLPSSSSRSLDSSSRSEKDRSLERERIGLGNYP 865  
DB 121 AATCEPASPAALSIFLSPFLOGAALARTPTPRG--GKEAEKDTSGE-----166  
767 ATENPTGSKRGLQISTAAQIAKMBEVSIAHTSQEDRSSGSTTELHLCHVTDERNALRS 925  
DB 167 ---AAVAAKAKALAVARIDQLVEDISALHTSSDDSFSLSS-----GDPGQEAPE 216  
768 SAAHTHS--NTYNTFKSENRTCSMPYAKLEYKRSNDSLSNVSNDGYGKGGQMKPSI 983  
DB 217 GRAQSCPCRGEGREAGRAHPLRLKAHASLNSDLSGASDGYCPREHM-----272  
769 ESYSEDESFKSCYQYPADLAHKIHSANHMDDNGELDTPIYSLKYSDEQLNSGRQSP 1043  
DB 273 -----LPCPLA-----ALASRRDP 287  
770 SONERWAPKHHIEDEIKQSEORSNQSTTYPVYTESTDDKILKQPHFGQECVSPYR 1103  
DB 288 -----RCQQRPP 294  
771 SRGANGSETNRVSGNHGINVQSLSQEDDYEKKPTNYSERYSEEEHEBERPTNYS 1163  
DB 295 SR-----LDLDLPGCAEPPAREATSADARVT---322  
772 IKYNEKRHVDDPIIDYSLKVATDIPSSQKQSFBSKSSQSSQSKTHEMSSSENTSTPSS 1223  
DB 323 IKLSPTYQHV-----PILLEGASRAGAEPFLAGFI-----SP 353  
773 NAKRONOLHFSSAQSRSGQPKAATCKVSSINQETITQYCEVDTPTCFRCSLSLSLSA 1283  
DB 354 GARKQAWL---PADHLKSVKEKLAAPL-SVASKALQKLAQGGPLSLRCSLSLSLSA 409  
774 EDEIGNQTTQEDASANTLOIAIEIKGIGTRSAE-----DPVSRVPAVSQHPRTKSR 1336  
DB 410 -GRPGPSEGGLDSDSSLEGL---EAGPSEALDSTWRAPGATSLPVAIPAPR---R 461  
775 LOGSSLSSEARHKAVERFPFGSAKSPSKGATPKSPPEHYVOETPLMFRCSTSVSLDSF 1396  
DB 462 NRGRLGVEDA-----TPSSSENIVQETPLVLGRCSVSSLSGSF 501  
776 ESASIASVQSEPCSGMVGIIISPSDLPSDPTGPMPPRSKTPP---PPQTAQTKREVEK 1454  
DB 502 ESPSIASSIPSEPCSGGSGTISPSSELPSDPTGPMPPRSKTPPPLAPAPG-----PP 554  
777 NKAPTAERESGPKQAANVAORVQLP---PADTILLHFAETSTPDGFCSSSSLSALSJ 1511  
DB 555 EATQFSLQWSEYKRFIDTADRCERCLPSELDAQS-VRFTEVKPDENFSCASSLSAL 613  
778 DEFFIQKQDELIMPPVQENDNGNETESEQPKESNENQKEAKETIDSEKLLDSDDDDD 1571  
DB 614 HEHVQQDVELRLPLSACPCEGGGAGGLHFAGHRRREEGPAPTGRPR---GAADQE 669  
779 IEILEECIISAMTKSRKKGKPAQATKALPPPVVARKPQLPVYKLLPONLQKQKVS 1631  
DB 670 LELLRECLGAAPVARLK-----VASALVP--GRRALFPVYMLVPAPAPAQ-----714  
780 FTPGDDMPVYCVGEGTPIFSTATISDLTIESFPNLAAGCVRGQAQSGEFKRDITP 1691

DB 715 ----EDDSCDTSAGCTPVNFSSAASLSDETLLQPPRDQFGGPAGR-----QRPT--759  
QY 1692 TEGRSTDEAOGKTSSVTIPELDDNKAEGDILAEICINSAMPKGSKSHKPPFRVKIMDOVQ 1751  
DB 760 --GRPTSARQA-----MGRHRK-----774  
QY 1752 QASASSAPNNQLDGGK-----KKPTSPVKPIQNTTEYRTRVRKNADSKNNLMAE 1802  
DB 775 --AGGAGRAEQSRGAGKNRAGLEPLGRPPSAPA-----807  
QY 1803 RVFSNKKSKQNLKNNKSKDFNDKLPNNEDRVRSFAPDSPHHYTPTEGTPYCFSRNDSL 1862  
DB 808 ----DKDGRPG-----RTRGDGALQSLCLTPTTEAVYCFYGNDS-844  
QY 1863 SSLDDDDDDVLSREKAE LRKAKENKESAEKVTSHTLTSNQQSANKTQAIKQPINRGO 1922  
DB 845 -----DEPPAAAPTTH-----RTSALPR-AFTREER 871  
QY 1923 PKPILQKSTFPQSKOIPDRGAATDEKLO-NFAIENTPVCFSHNSLSLSLSDIDQNNN 1981  
DB 872 P-----QGRKEAPAPSKAAP--AAPPARTQPSLIADETPCYSLSSSASSLS-----917  
QY 1982 KENEPIKETEPDSDGSEPSKQASGYAPKSPHVEDTPVCFSRNSSLSSLSIDSSEDLLOE 2041  
DB 918 -EPEP--SEPPAVHPRGREFAVT-----KDPGCGGRDSSPSP--RAABEELQOR 961  
QY 2042 CISSAMPKXKXKX-----SRLKGDNEKHSFRNMGGILGEDTLIDL 2080  
DB 962 CISSALPRRPPVSGLRKRRKPRATRLDERPAEGSRERGEAAAGSDR-----1007  
QY 2081 KDIQEPDSEHGLSPDSENFDMKATOBGANSIVSLHQAARAAACLSRQASDSDSILSLS 2140  
DB 1008 -----ASLDSVWEVRAIOEGANSIVTLHQAARAA--TREASSEDLSLFSVS 1052  
QY 2141 GISLGS-----PFHLTPQOEEKPFSTNKGPRILKDGKSTLETTKIESKIGKGGKVKYK 2196  
DB 1053 GLSVGSTLQPKHKKRGAEGENGSAARP-----EKRGAAVKTSGSPSPAGPEK--1103  
QY 2197 SLITGKVRNSIEISGQKKQKLOANMPSISRGRTMIHI-----FGVRNNSSSSTS 2244  
DB 1104 -----PRGTQKTTGVPALVRGRVIVYVSPAPRAQPKGTGEPATPRKVA 1149  
QY 2245 PVSXKGPPLKTPA--SKSPSEGTATTS--PRGAKPSVKSELSPVARQTSOIGGSSKAPSR 2301  
DB 1150 P-----PCLAQAPAPAKVPSPGQORSLSRHPAKTSBLATLSQPPRSATPPARLAKTPSS 1204  
QY 2302 SGSRDSTPSRPAQOPLSR---PIOSPGRNSISPGRNGISPPNKLSQLPRTSSPSTASTKS 2358  
DB 1205 SSSQTS-----PASQPLPRKRPVVTQAA-----GALFPGGASPVPKTPARTLLAKQH 1251  
QY 2359 SGSGKMSVTSRGRQMSQNLTKOTGLSKNASSIPRSESASKGLNQMNGNKA-NKKVELS 2417  
DB 1252 K-----TQSRPVRLPFMORPARP-GPPPLARAVP--EPGPRGAGTAGCARGRUGLV 1303  
QY 2418 RMSSTKSGSSSDSRSPVLVVRQSTFIKEAPSTPLRKLSEASFEISLSPSSPASPTRS 2477  
DB 1304 RVASALSSGSES--SDRSGRFRQLTFIKE--SPGLRRRRSELSSAESAAAPAGASPRRG 1359  
QY 2478 QAQTPVLSPLPDMSLSTHSSVQ-----AGWRKLPNLSPTTYNDGRKAKRHDA 2529  
DB 1360 R-----PALPVLFCSSRCELRAPAPORPPAPARPS-----PGERPAR-----1404  
QY 2530 RSHSESRLPPINRSGTWKREHSHSSSLPRVSTWRTTSSSSTILGASSSSSSSKAKSEDE 2589  
DB 1405 RTTSESRLPVP-RAPARPETVKRYASLPHISVARRPDGAVPAPASADAARSSDGE 1463  
QY 2590 KXVNSISGTQSKENQVSAKG-TWRKIKENEFF-----SPINSTSQTVSSGATNGASKTL 2643  
DB 1464 RPL-----PRVAAPGTTWRRIRDEVDVPHILNSTLPATLPLRGSTPBDAPA---1509  
QY 2644 IYQVAPAVSKTEDVWVRIEDCPINPRSGRSPGNT--PPVIDSVSEKANPNIKSDKNQ 2701  
DB 1510 ----GPPPKTSDAVVQTEEVAAPKTNSTSPSILETEPFGAPAGQ-----1552

Search completed: August 25, 2004, 17:52:37  
Job time : 216.5 secs

**This Page Blank (uspto)**

nCore version 5.1.6  
1993 - 2004 CompuGen Ltd.

ing sw model

, 17:08:42 ; Search time 37.5 Seconds  
(without alignments)  
3913.937 Million cell updates/sec

.2

FEALKMN.....ESSGTQSPKRGHSYLVTSV 2843

ixpt 0.5

.625971 residues

i chosen parameters: 389414

100

0%

100%

45 summaries

i AA: \*  
xdata/2/iaa/5A\_COMB.pep: \*  
xdata/2/iaa/5B\_COMB.pep: \*  
xdata/2/iaa/6A\_COMB.pep: \*  
xdata/2/iaa/6B\_COMB.pep: \*  
xdata/2/iaa/PCTUS\_COMB.pep: \*  
xdata/2/iaa/backfiles1.pep: \*

of results predicted by chance to have a  
ial to the score of the result being printed,  
is of the total score distribution.

SUMMARIES

3	ID	Description
1	US-07-741-940-2	Sequence 2, Appli
1	US-08-289-548A-2	Sequence 2, Appli
1	US-08-452-654-2	Sequence 2, Appli
2	US-08-370-235A-2	Sequence 2, Appli
1	US-08-449-731-2	Sequence 2, Appli
1	US-08-452-655B-2	Sequence 2, Appli
1	US-08-452-655B-7	Sequence 7, Appli
3	US-08-450-582-2	Sequence 2, Appli
2	US-08-821-355A-7	Sequence 7, Appli
2	US-09-003-687A-7	Sequence 7, Appli
3	US-09-136-605-7	Sequence 7, Appli
1	US-07-741-940-7	Sequence 7, Appli
1	US-08-289-548A-7	Sequence 7, Appli
1	US-08-452-654-7	Sequence 7, Appli
4	US-08-449-731-7	Sequence 7, Appli
2	US-08-630-822A-100	Sequence 100, App
2	US-09-005-069-100	Sequence 100, App
4	US-09-171-156A-49	Sequence 49, Appl
4	US-09-004-730A-49	Sequence 49, Appl
4	US-08-981-799A-49	Sequence 49, Appl
4	US-09-134-001C-4463	Sequence 4463, Ap
4	US-09-134-001C-3159	Sequence 3159, Ap
4	US-09-134-001C-5080	Sequence 5080, Ap
1	US-08-353-700-1	Sequence 1, Appli
5	PCT-US55-16216-1	Sequence 1, Appli
4	US-08-714-741-32	Sequence 32, Appli

28	371.5	2.5	1581	3	US-09-110-517-2	Sequence 2, Appli
29	371.5	2.5	3256	4	US-09-919-172-98	Sequence 98, Appli
30	371.5	2.5	3256	4	US-09-976-594-22	Sequence 22, Appli
31	371.5	2.5	3969	3	US-08-061-376-5	Sequence 5, Appli
32	371	2.5	2468	4	US-09-976-594-726	Sequence 726, App
33	355	2.4	2662	4	US-09-595-684B-31	Sequence 31, Appli
34	336	2.3	2954	4	US-09-150-867-1	Sequence 1, Appli
35	334.5	2.3	2185	4	US-09-854-856-36	Sequence 36, Appli
36	334.5	2.3	2245	4	US-09-854-856-4	Sequence 4, Appli
37	334.5	2.3	2332	4	US-09-854-856-34	Sequence 34, Appli
38	334.5	2.3	2382	4	US-09-854-856-2	Sequence 2, Appli
39	326.5	2.2	2482	1	US-08-328-254-6	Sequence 6, Appli
40	321	2.2	2157	4	US-09-854-856-52	Sequence 52, Appli
41	321	2.2	2217	4	US-09-854-856-20	Sequence 20, Appli
42	321	2.2	2294	4	US-09-854-856-50	Sequence 50, Appli
43	321	2.2	2354	4	US-09-854-856-18	Sequence 18, Appli
44	315.5	2.2	1596	4	US-08-978-277A-4	Sequence 4, Appli
45	313.5	2.2	1863	1	US-08-425-061-16	Sequence 16, Appli

ALIGNMENTS

RESULT 1  
US-07-741-940-2  
; Sequence 2, Application US/07741940  
; Patent No. S352775  
; GENERAL INFORMATION:  
; APPLICANT: ALBERTSEN, HANS  
; APPLICANT: ANAND, RAKESH  
; APPLICANT: CARLSON, MARY  
; APPLICANT: GORDEN, JOANNA  
; APPLICANT: HEDGE, PHILIP J.  
; APPLICANT: JOSLYN, GEOFF  
; APPLICANT: KINZLER, KENNETH  
; APPLICANT: MARKHAM, ALEXANDER F.  
; APPLICANT: NAKAMURA, YUSUKE  
; APPLICANT: THLIVERIS, ANDREW  
; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC  
; TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS  
; NUMBER OF SEQUENCES: 94  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Banner, Birch, McKie & Beckett  
; STREET: 1001 G Street, NW  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20001-4598  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07741,940  
; FILING DATE: 19920109  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kagan, Sarah A.  
; REGISTRATION NUMBER: 32,141  
; REFERENCE/DOCKET NUMBER: 1107.035574  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-508-9100  
; TELEFAX: 202-508-9299  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2843 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-741-940-2  
Query Match 100.0%; Score 14575; DB 1; Length 2843;

Best Local Similarity 100.0%; Pred. No. 0;			
Matches 2843; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	MAAASYDOLLKQVAKLVENSLRQBLDQNSNHLTKLETSANMKEVVKLOQSIEDAM	60
Db	1	MAAASYDOLLKQVAKLVENSLRQBLDQNSNHLTKLETSANMKEVVKLOQSIEDAM	60
Qy	61	ASSGQIDLERLKEINLDSNFPVGVKLRSMKLSRYSYSGRSYSSRSGECSPPVMSGFP	120
Db	61	ASSGQIDLERLKEINLDSNFPVGVKLRSMKLSRYSYSGRSYSSRSGECSPPVMSGFP	120
Qy	121	RGFVNGSRSTGYLEERKERSLLADLKEEKEKDWYQAQLQNLTKRIDSLPLTENFSL	180
Db	121	RGFVNGSRSTGYLEERKERSLLADLKEEKEKDWYQAQLQNLTKRIDSLPLTENFSL	180
Qy	181	QTDLTRRQLEYEARQIRVAMEEQIGTQDMKRAORRIARIQOIEKDKILAIRQLQSQAT	240
Db	181	QTDLTRRQLEYEARQIRVAMEEQIGTQDMKRAORRIARIQOIEKDKILAIRQLQSQAT	240
Qy	241	EAERSQNKHETGSHDAERQEGGVGEINMATSNGQGGSTTRMDHETASVLSSTHSA	300
Db	241	EAERSQNKHETGSHDAERQEGGVGEINMATSNGQGGSTTRMDHETASVLSSTHSA	300
Qy	301	PRRLTSHLGTVMVYSLMLGTHDXDDMSRTLLAMSSODSCISMRQSGCLPLLIQLL	360
Db	301	PRRLTSHLGTVMVYSLMLGTHDXDDMSRTLLAMSSODSCISMRQSGCLPLLIQLL	360
Qy	361	HGNDKDSVLLNGSRGSEARASAAALHNIHSQPDQKGRREIRVLHLLQIRAYCETC	420
Db	361	HGNDKDSVLLNGSRGSEARASAAALHNIHSQPDQKGRREIRVLHLLQIRAYCETC	420
Qy	421	WEOQEAHPGMDQKNPMPAPVEHQICPAVCVLMKLSFDEHRHAMNELGGLQAIABLLQ	480
Db	421	WEOQEAHPGMDQKNPMPAPVEHQICPAVCVLMKLSFDEHRHAMNELGGLQAIABLLQ	480
Qy	481	VDCMYGLTNDHYITLRRVAGMALTNLTGCDVANKATLCSMKGCMFALVAQLKSEEDL	540
Db	481	VDCMYGLTNDHYITLRRVAGMALTNLTGCDVANKATLCSMKGCMFALVAQLKSEEDL	540
Qy	541	QQVIASVLRNLWSRADVNSKTLREVGSVKALMECALEVKKESTKLSVLSALNLSAHT	600
Db	541	QQVIASVLRNLWSRADVNSKTLREVGSVKALMECALEVKKESTKLSVLSALNLSAHT	600
Qy	601	ENKADICAVDGCALAFVGLTYRSQTNLTALIESGGGILRNVSLLIATNEDHRIQLRENN	660
Db	601	ENKADICAVDGCALAFVGLTYRSQTNLTALIESGGGILRNVSLLIATNEDHRIQLRENN	660
Qy	661	CLQTLQLKSHSLTIYSNACGLTNLNSARNPKDQEAALWDMGAVSMLKNIHSHKHMIAM	720
Db	661	CLQTLQLKSHSLTIYSNACGLTNLNSARNPKDQEAALWDMGAVSMLKNIHSHKHMIAM	720
Qy	721	GSAALRNLANRPAPKADANIKSPGSLPSLHVROKALEAELDACHLSETFDNTDNL	780
Db	721	GSAALRNLANRPAPKADANIKSPGSLPSLHVROKALEAELDACHLSETFDNTDNL	780
Qy	781	PKASHRSKORHKSQLYGDIYFDTNRHDDNSDNFNTGNMTVLPYLNNTIVLPSSSSSRGS	840
Db	781	PKASHRSKORHKSQLYGDIYFDTNRHDDNSDNFNTGNMTVLPYLNNTIVLPSSSSSRGS	840
Qy	841	LDSSRSKORSLERERIGIGVNHYPATENGTSSKRGLOISTTAAQIAKWMESVAIHTS	900
Db	841	LDSSRSKORSLERERIGIGVNHYPATENGTSSKRGLOISTTAAQIAKWMESVAIHTS	900
Qy	901	QEDRSSGSTTDLHCVTDERNALRRSSAAHSTHNTYNTFKSENENRITCSMPYAKLEYKRS	960
Db	901	QEDRSSGSTTDLHCVTDERNALRRSSAAHSTHNTYNTFKSENENRITCSMPYAKLEYKRS	960
Qy	961	NDSLNSVSSNDGYKRGOMKPSIESYEDDESFKCSYGQYPADLAHKIHSANHMDNDGE	1020
Db	961	NDSLNSVSSNDGYKRGOMKPSIESYEDDESFKCSYGQYPADLAHKIHSANHMDNDGE	1020
Qy	1021	LDTPIYSLKYSDQLNSGRQSPQNERWAPKPHIIEDEIKQSEQRQSRNQSTTVPVYTE	1080
Db	1021	LDTPIYSLKYSDQLNSGRQSPQNERWAPKPHIIEDEIKQSEQRQSRNQSTTVPVYTE	1080

QY 2161 SNKGPRILKPGKSTLTETKKIESSEKIGKGGKVKVYKSLITGKVRNSEISGOMKOPLOAN 2220  
Db 2161 SNKGPRILKPGKSTLTETKKIESSEKIGKGGKVKVYKSLITGKVRNSEISGOMKOPLOAN 2220  
QY 2221 MPSISGRMTMIHPGVNRSSSTSPVSKGPKLTPASKSPSEGTATTSPRGAKPSVKS 2280  
Db 2221 MPSISGRMTMIHPGVNRSSSTSPVSKGPKLTPASKSPSEGTATTSPRGAKPSVKS 2280  
QY 2281 ELSVARQTSQIGGSKAPSRGSRDSTPSRPAQOPLSRPIOSPGRNISTPGRNGISPPN 2340  
Db 2281 ELSVARQTSQIGGSKAPSRGSRDSTPSRPAQOPLSRPIOSPGRNISTPGRNGISPPN 2340  
QY 2341 KLSQLPRTSSPTASTKSSGSGKMSYTSPPGRQMSQONLTQTGLSKNASSIPRSESASKG 2400  
Db 2341 KLSQLPRTSSPTASTKSSGSGKMSYTSPPGRQMSQONLTQTGLSKNASSIPRSESASKG 2400  
QY 2401 LQNMNGNGANKKVELSRMSSSTKSSGSDRSPVLRQSTFIKEAPSPILRRKLEBSA 2460  
Db 2401 LQNMNGNGANKKVELSRMSSSTKSSGSDRSPVLRQSTFIKEAPSPILRRKLEBSA 2460  
QY 2461 SPESISPSRRPASPTRSQAOTPVLSPLPDMSLSTHSSVQAGGWRKLPNLSPTIEYNDG 2520  
Db 2461 SPESISPSRRPASPTRSQAOTPVLSPLPDMSLSTHSSVQAGGWRKLPNLSPTIEYNDG 2520  
QY 2521 RPAKHDIARSHSESPSRIPINRSGTWKREHKHSSSLPRVSTWRTGSSSILSASSES 2580  
Db 2521 RPAKHDIARSHSESPSRIPINRSGTWKREHKHSSSLPRVSTWRTGSSSILSASSES 2580  
QY 2581 SEKAKSEDEKHVNSISGTTQSKENOVSAKGTWRKIKENEFSTNTSTQTVSSGATNGAES 2640  
Db 2581 SEKAKSEDEKHVNSISGTTQSKENOVSAKGTWRKIKENEFSTNTSTQTVSSGATNGAES 2640  
QY 2641 KTLIQMAPAVAKTEDVWRIEDCPINPRSGRSTGNTPTPVVIDSVSEKAMPNIKSDKN 2700  
Db 2641 KTLIQMAPAVAKTEDVWRIEDCPINPRSGRSTGNTPTPVVIDSVSEKAMPNIKSDKN 2700  
QY 2701 QAKQNVGSGVPMRTVGLNRLTSTFIQVADPQKTEIKPGONNPVPSVETNESPIVERT 2760  
Db 2701 QAKQNVGSGVPMRTVGLNRLTSTFIQVADPQKTEIKPGONNPVPSVETNESPIVERT 2760  
QY 2761 PFSSSSSKHSGSPSGTVAARVTPFNPNPSRKSSADSTARSPSQIPTPVNNTKKRDSKT 2820  
Db 2761 PFSSSSSKHSGSPSGTVAARVTPFNPNPSRKSSADSTARSPSQIPTPVNNTKKRDSKT 2820  
QY 2821 DSTESSGTQSPKRGSGSVLVTGV 2843  
Db 2821 DSTESSGTQSPKRGSGSVLVTGV 2843

RESULT 2  
US-08-289-548A-2  
; Sequence 2, Application US/08289548A  
; Patent No. 5648212  
; GENERAL INFORMATION:  
; APPLICANT: ALBERTSEN, HANS  
; APPLICANT: ANAND, RAKESH  
; APPLICANT: CARLSON, MARY  
; APPLICANT: GRODEN, JOANNA  
; APPLICANT: HEDGE, PHILIP J.  
; APPLICANT: JOSLYN, GEOFF  
; APPLICANT: KINZLER, KENNETH  
; APPLICANT: MARKHAM, ALEXANDER F.  
; APPLICANT: NAKAMURA, YUSUKE  
; APPLICANT: TILVERIS, ANDREW  
; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC  
; NUMBER OF SEQUENCES: 102  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Banner & Allegretti, LTD  
; STREET: 1001 G Street, NW  
; CITY: Washington  
; STATE: D.C.

COUNTRY: USA  
ZIP: 20001-4598  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/289,548A  
FILING DATE: 12-AUG-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Kagan, Sarah A.  
REGISTRATION NUMBER: 32,141  
REFERENCE/DOCKET NUMBER: 1107.46943  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-508-9100  
TELEFAX: 202-508-9239  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2843 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-289-548A-2

Query Match 100.0%; Score 14575; DB 1; Length 2843;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2843; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MAAASYDOLLKQVEALKWENSRLQELNDSNHLTKLETSANMKVYLKQCSIEDAM 60  
Db 1 MAAASYDOLLKQVEALKWENSRLQELNDSNHLTKLETSANMKVYLKQCSIEDAM 60  
QY 61 ASSQIDLLERLKEINLDSSNPFQVKLRKMSLRYSYSGREGSVSSRSGECPVPMGSFPR 120  
Db 61 ASSQIDLLERLKEINLDSSNPFQVKLRKMSLRYSYSGREGSVSSRSGECPVPMGSFPR 120  
QY 121 RGFVNGSRESTGYLEELKERSLLADLDKEEKEKQWYQAOLQNLTKRISLPIETENFSL 180  
Db 121 RGFVNGSRESTGYLEELKERSLLADLDKEEKEKQWYQAOLQNLTKRISLPIETENFSL 180  
QY 181 QTLTTRQLYEAREQIRVAMEEQLGTQDMKEKAQRRIARIQIIEKDILRIQLQSOAT 240  
Db 181 QTLTTRQLYEAREQIRVAMEEQLGTQDMKEKAQRRIARIQIIEKDILRIQLQSOAT 240  
QY 241 EAERSQNKHETGSHDAERQNEGGVGEINMATSGNGQSGTTRMDHETASVLSSTHSA 300  
Db 241 EAERSQNKHETGSHDAERQNEGGVGEINMATSGNGQSGTTRMDHETASVLSSTHSA 300  
QY 301 PRRLTSHLGTKEVMYSLLSMLGTHDKDMSRTLLAMSSODSCI SMRQSGCLPLLIQLL 360  
Db 301 PRRLTSHLGTKEVMYSLLSMLGTHDKDMSRTLLAMSSODSCI SMRQSGCLPLLIQLL 360  
QY 361 HGNDKDSVLLGNSRGSKAEARASAAALHNIHSGQDDPKRGRREIRVHLLEQIRAYCETC 420  
Db 361 HGNDKDSVLLGNSRGSKAEARASAAALHNIHSGQDDPKRGRREIRVHLLEQIRAYCETC 420  
QY 421 WEQEAHEPGMDQDKNPMAPAVEHQICPAVCVLMKLSFDEEHRHANNELGGLQAIALLQ 480  
Db 421 WEQEAHEPGMDQDKNPMAPAVEHQICPAVCVLMKLSFDEEHRHANNELGGLQAIALLQ 480  
QY 481 VDCEMYGLTNDHYSITLRRYAGVALTNLTFDGVANKATLCSMGKCMRALVAQKSESEDL 540  
Db 481 VDCEMYGLTNDHYSITLRRYAGVALTNLTFDGVANKATLCSMGKCMRALVAQKSESEDL 540  
QY 541 QQVIASVRLNLSWRADVNSKKTIREVGSVKALMECALEVKKESTLKSVALNLSAHC 600  
Db 541 QQVIASVRLNLSWRADVNSKKTIREVGSVKALMECALEVKKESTLKSVALNLSAHC 600  
QY 601 ENKADICAVDGAFLAVGLTITYSQNTNLAIESGGILRNVSLLIATNEDHQILRENN 660  
Db 601 ENKADICAVDGAFLAVGLTITYSQNTNLAIESGGILRNVSLLIATNEDHQILRENN 660

QY 661 CLQTLLOHLKSHSLTIVSNACGLTWNLSARNPKDQOALWDMGAVSMKLNLIHSHKHQIAM 720  
 DB 661 CLQTLLOHLKSHSLTIVSNACGLTWNLSARNPKDQOALWDMGAVSMKLNLIHSHKHQIAM 720  
 QY 721 GSAALANLANPAKYKANDANIMSPGSLPSLHVKQKALAEALDAHLSEFDNIDNLS 780  
 DB 721 GSAALANLANPAKYKANDANIMSPGSLPSLHVKQKALAEALDAHLSEFDNIDNLS 780  
 QY 781 PKASHRSKQKHQSLYGDYVDFNHRDDNRDNFNNTGNMTVLSPYLVNTVLPPSSSSRGS 840  
 DB 781 PKASHRSKQKHQSLYGDYVDFNHRDDNRDNFNNTGNMTVLSPYLVNTVLPPSSSSRGS 840  
 QY 841 LQSSRSKDLSEERIGIGLVNHPATENPGTSSKRGLOLSTTAAOIAKYMEEVSAIHTS 900  
 DB 841 LQSSRSKDLSEERIGIGLVNHPATENPGTSSKRGLOLSTTAAOIAKYMEEVSAIHTS 900  
 QY 901 QEDRSSGTTTELHCVTDERNALRASSAAHTSNFTKSENNRCTCMPYAKLEYKRSS 960  
 DB 901 QEDRSSGTTTELHCVTDERNALRASSAAHTSNFTKSENNRCTCMPYAKLEYKRSS 960  
 QY 961 NDSLSNVSNDGKRGOMKPSIBSYSEDDSKFCYGOYPADLAHKIHSANMDNDGE 1020  
 DB 961 NDSLSNVSNDGKRGOMKPSIBSYSEDDSKFCYGOYPADLAHKIHSANMDNDGE 1020  
 QY 1021 LDTPINTSLKYSDBQLNSGRQSPQNERWARPKHIIIEDEIKQSEQRQSRNQSTTYPVYTE 1080  
 DB 1021 LDTPINTSLKYSDBQLNSGRQSPQNERWARPKHIIIEDEIKQSEQRQSRNQSTTYPVYTE 1080  
 QY 1081 STDDKHLKQFHFQOQECVSPYRSGANGSETNRVGSNHGQINQVOSLCOEDDYDDKP 1140  
 DB 1081 STDDKHLKQFHFQOQECVSPYRSGANGSETNRVGSNHGQINQVOSLCOEDDYDDKP 1140  
 QY 1141 TNYERYSEEBQHEBERPTNYSIKYNEKXHVDPIDYSLKYATDIPSSQKQSFQSKS 1200  
 DB 1141 TNYERYSEEBQHEBERPTNYSIKYNEKXHVDPIDYSLKYATDIPSSQKQSFQSKS 1200  
 QY 1201 SSGOSSKTEHMSSENSTSPSSNAKQNLHPSSAQSRSQGPQKAACTKVSSINQETIQ 1260  
 DB 1201 SSGOSSKTEHMSSENSTSPSSNAKQNLHPSSAQSRSQGPQKAACTKVSSINQETIQ 1260  
 QY 1261 TYCVEDTPICFRCSLSLSAASDEIGCNOTTOEADSANLQIAELKIGKIGTRSAEDPV 1320  
 DB 1261 TYCVEDTPICFRCSLSLSAASDEIGCNOTTOEADSANLQIAELKIGKIGTRSAEDPV 1320  
 QY 1321 SEVPAVSQHPRTKSRLOGSSLSSESAPHKAVEPPSGAKSPSKSGAQTPKSPPEHYVQET 1380  
 DB 1321 SEVPAVSQHPRTKSRLOGSSLSSESAPHKAVEPPSGAKSPSKSGAQTPKSPPEHYVQET 1380  
 QY 1381 PLMFSTRCTSVSLDSFESRSITASSVQSEPCSGMGVGIISPSDLDPSPQGTWMPRSRKTTP 1440  
 DB 1381 PLMFSTRCTSVSLDSFESRSITASSVQSEPCSGMGVGIISPSDLDPSPQGTWMPRSRKTTP 1440  
 QY 1441 PPPQTAQTKREVVPKNKAPTAEKRESGPKQAANVAQVVOVLPDADTLLHFATESTPDGF 1500  
 DB 1441 PPPQTAQTKREVVPKNKAPTAEKRESGPKQAANVAQVVOVLPDADTLLHFATESTPDGF 1500  
 QY 1501 SCSSLSLSLDEBPIQKDELRITMPPVQENDNGNETESQPKESNENQKEAEKTIIDSE 1560  
 DB 1501 SCSSLSLSLDEBPIQKDELRITMPPVQENDNGNETESQPKESNENQKEAEKTIIDSE 1560  
 QY 1561 KDLDDSDDDDEIELEECIIISAMPTKSRKGGKPAQATASKLPPPVAPKPSQLPVYKLLPS 1620  
 DB 1561 KDLDDSDDDDEIELEECIIISAMPTKSRKGGKPAQATASKLPPPVAPKPSQLPVYKLLPS 1620  
 QY 1621 QNRLOPQKHVSFTPGDDMPRVYVEGTPINFTSTATSLSDLTIESPPNELLAGEVGRGAQ 1680  
 DB 1621 QNRLOPQKHVSFTPGDDMPRVYVEGTPINFTSTATSLSDLTIESPPNELLAGEVGRGAQ 1680  
 QY 1681 SGEFEKRTIIPTEGRSTDEAOGKTSVTIPELDDNKAEEGDILAEACINSAMPKKGSKHP 1740  
 DB 1681 SGEFEKRTIIPTEGRSTDEAOGKTSVTIPELDDNKAEEGDILAEACINSAMPKKGSKHP 1740

QY 1741 FRVKKIMQVOQASASSAPNKNQDLGKKKKPTSPVKPIQNTTEYRTRVRKNADSKNNLN 1800  
 DB 1741 FRVKKIMQVOQASASSAPNKNQDLGKKKKPTSPVKPIQNTTEYRTRVRKNADSKNNLN 1800  
 QY 1801 AERVFSDKDKKQNLKONSKDFNDKLPNNEDRVRGSFAFDSPHHYTPIEGTPYCFGRND 1860  
 DB 1801 AERVFSDKDKKQNLKONSKDFNDKLPNNEDRVRGSFAFDSPHHYTPIEGTPYCFGRND 1860  
 QY 1861 SLSLSDFDDDDDVLSREKAEARAKENKESEAKVTSHTELTSSQOSANKTQAIAKQPINR 1920  
 DB 1861 SLSLSDFDDDDDVLSREKAEARAKENKESEAKVTSHTELTSSQOSANKTQAIAKQPINR 1920  
 QY 1921 GQPKPILQKQSTFPQSSKQIPDRGAATDEKLQNFATIENTPVCFSHNSLSLSLSDIDQENN 1980  
 DB 1921 GQPKPILQKQSTFPQSSKQIPDRGAATDEKLQNFATIENTPVCFSHNSLSLSLSDIDQENN 1980  
 QY 1981 NKENEPKETETPPDSQCEPSPKQASGYAPKSFHVEDTPVCFSRNSSLSLSLSDIDEDDLQ 2040  
 DB 1981 NKENEPKETETPPDSQCEPSPKQASGYAPKSFHVEDTPVCFSRNSSLSLSLSDIDEDDLQ 2040  
 QY 2041 ECISAMPKPKKPKSRLKGDNEKHSFRNMGGILGEDLTLDLKDIOQRPSEHGLSPDSNFD 2100  
 DB 2041 ECISAMPKPKKPKSRLKGDNEKHSFRNMGGILGEDLTLDLKDIOQRPSEHGLSPDSNFD 2100  
 QY 2101 WKATOEGANSIVSSLHQAAAAACLSRQASSDSDSLKSGISLGSPPHLPDQOEBKDET 2160  
 DB 2101 WKATOEGANSIVSSLHQAAAAACLSRQASSDSDSLKSGISLGSPPHLPDQOEBKDET 2160  
 QY 2161 SNKGPRILLKPEKSTLETYKIESKESKGIKGGKYVYKSLITGKVRNSNIEISQMKQLOAN 2220  
 DB 2161 SNKGPRILLKPEKSTLETYKIESKESKGIKGGKYVYKSLITGKVRNSNIEISQMKQLOAN 2220  
 QY 2221 MPSISRGTMTHIPGVNRSSSTSPVSKGKGLPPLTPASKSPSEGOATTSPRGAKPSVKS 2280  
 DB 2221 MPSISRGTMTHIPGVNRSSSTSPVSKGKGLPPLTPASKSPSEGOATTSPRGAKPSVKS 2280  
 QY 2281 ELSVPARTSOIIGSSKAPSSGSRDSTSPRPAQPLSRPIQSPGRNSISFGRNGISPPN 2340  
 DB 2281 ELSVPARTSOIIGSSKAPSSGSRDSTSPRPAQPLSRPIQSPGRNSISFGRNGISPPN 2340  
 QY 2341 KLSQPLRTSSPSTASTKSSGSKXSYTSPGQMSQOQLTKOTGLSKXASSTIPRESASKG 2400  
 DB 2341 KLSQPLRTSSPSTASTKSSGSKXSYTSPGQMSQOQLTKOTGLSKXASSTIPRESASKG 2400  
 QY 2401 LQNMNGANGANKVELSRWSTKSSGSDSRSEPRVLVRQSTTFKEAPSPTRLKLEESA 2460  
 DB 2401 LQNMNGANGANKVELSRWSTKSSGSDSRSEPRVLVRQSTTFKEAPSPTRLKLEESA 2460  
 QY 2461 SPESLSPSSRPASPTRSQAQTPVLSPLDMSLSTHSSVQAGWKRKUPPNLSPTIENDG 2520  
 DB 2461 SPESLSPSSRPASPTRSQAQTPVLSPLDMSLSTHSSVQAGWKRKUPPNLSPTIENDG 2520  
 QY 2521 RPAKHDIARSHESPSRLPINRSGTWKREHSKHSSSLPRVSTWERTGSSSSILSASSES 2580  
 DB 2521 RPAKHDIARSHESPSRLPINRSGTWKREHSKHSSSLPRVSTWERTGSSSSILSASSES 2580  
 QY 2581 SEKASXSEKHNYSISGTTKQSKENQVSAKGTWRKIKENEFSTNSTQTSVSSGATNGAES 2640  
 DB 2581 SEKASXSEKHNYSISGTTKQSKENQVSAKGTWRKIKENEFSTNSTQTSVSSGATNGAES 2640  
 QY 2641 KTLIYQMAPAVSKTEDVWVRIDECPIINPRSGRSPGTGTPVIDSVSEKAPNLIKSDKN 2700  
 DB 2641 KTLIYQMAPAVSKTEDVWVRIDECPIINPRSGRSPGTGTPVIDSVSEKAPNLIKSDKN 2700  
 QY 2701 QAKQNVGNGSVPMRTVGLNRLTFSIQVADAPQCKTEIKPGQNNPVPVSETNESPIVERT 2760  
 DB 2701 QAKQNVGNGSVPMRTVGLNRLTFSIQVADAPQCKTEIKPGQNNPVPVSETNESPIVERT 2760  
 QY 2761 PFSSSSSSKHSSPSGTVAARVTPFNYPNPSPRKSSADSTSAQPSQIPTVNNNTKKRDSKT 2820  
 DB 2761 PFSSSSSSKHSSPSGTVAARVTPFNYPNPSPRKSSADSTSAQPSQIPTVNNNTKKRDSKT 2820  
 QY 2821 DSTFSSSGTQSPKRHSGSVLVTSTV 2843



```

Db      2821 DSTSSGQTSPRRHSGSYLVTSV 2843
|||||
RESULT 3
US-08-452-654-2
; Sequence 2, Application US/08452654
; Patent No. 5691454
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, HANS
; APPLICANT: ANAND, RAKESH
; APPLICANT: CARLSON, MARY
; APPLICANT: GRODEN, JOANNA
; APPLICANT: HEDGE, PHILIP J.
; APPLICANT: JOSLYN, GEOFF
; APPLICANT: KINZLER, KENNETH
; APPLICANT: MARKHAM, ALEXANDER F.
; APPLICANT: NAKAMURA, YUSUKE
; APPLICANT: THLIVERIS, ANDREW
; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC
; SITE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner, Birch, McKie & Beckett
; STREET: 1001 G Street, NW
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001-4598
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/452,654
; FILING DATE: 25-MAY-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/741,940
; FILING DATE: 08-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Kagan, Sarah A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 1107.035574
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9299
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2843 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-452-654-2
Query Match          100.0%; Score 14575; DB 1; Length 2843;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2843; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MAASVDQLLKQVKALKMNSNLRLQELLEDNSHLTKLETEASNKMKVLKQLGGSIEDAM 60
DB      1 MAASVDQLLKQVKALKMNSNLRLQELLEDNSHLTKLETEASNKMKVLKQLGGSIEDAM 60
QY      61 ASSGQIDLLRLKEINLDSSNFPQVKLRSKVRSLRYSVGSRREGSVSSRGSCSPVMGSPFR 120
DB      61 ASSGQIDLLRLKEINLDSSNFPQVKLRSKVRSLRYSVGSRREGSVSSRGSCSPVMGSPFR 120
QY      121 RGCVNGSRESTGYLLEELEKERSLLADLDKKEEKDWYYAQQLNTKRIDSPLTFENFSL 180
DB      121 RGCVNGSRESTGYLLEELEKERSLLADLDKKEEKDWYYAQQLNTKRIDSPLTFENFSL 180
QY      181 QTDLTRQLYEYARQIRVAMEEQLGTCCDQWKAQRRIARIQQIEKDILIRIQLLSQAT 240

```

Db 1361 TYCEDVTPICFSRCSLSLSLSSAEDEICGQNTQEQADSANTLOIAIEIKGKIGTRESAEDPV 1320  
Qy 1321 SEVPAVCHPRTKSRIQSSLSSESARHKAVERFPSCAKSPKSGAQTQPKSPPEHYVQET 1380  
Db 1321 SEVPAVCHPRTKSRIQSSLSSESARHKAVERFPSCAKSPKSGAQTQPKSPPEHYVQET 1380  
Qy 1381 PLMESRCTSVSLDSFESRSTASSVQSPSCGWSVGIISPSDLDPSPGQTMPSPRSKPTP 1440  
Db 1381 PLMESRCTSVSLDSFESRSTASSVQSPSCGWSVGIISPSDLDPSPGQTMPSPRSKPTP 1440  
Qy 1441 PPPQTAQTKREVKNKAPTAKEKRESGPKQAANAAVQVQLPDADTLLHFATESPDGF 1500  
Db 1441 PPPQTAQTKREVKNKAPTAKEKRESGPKQAANAAVQVQLPDADTLLHFATESPDGF 1500  
Qy 1501 SCSSLSALSILDEPFIQKVELRIMPVQENDNGNETSEQPKESNENQKEAEKIDSE 1560  
Db 1501 SCSSLSALSILDEPFIQKVELRIMPVQENDNGNETSEQPKESNENQKEAEKIDSE 1560  
Qy 1561 KDLDDDDDDDDIELEBECIIISAMPTKSRKGKPAQTAASKLPPPVARKPSQLPVYKLLPS 1620  
Db 1561 KDLDDDDDDDDIELEBECIIISAMPTKSRKGKPAQTAASKLPPPVARKPSQLPVYKLLPS 1620  
Qy 1621 QNRLOPKHVSFTPGDDMPVYCVGEGTPIINFSTATSLSDLTIESPPNELAAGSVRGGAQ 1680  
Db 1621 QNRLOPKHVSFTPGDDMPVYCVGEGTPIINFSTATSLSDLTIESPPNELAAGSVRGGAQ 1680  
Qy 1681 SGEFEKRDITPTEGRSDEAGQKTSVTIPELDDNKAEGDILAEICINAMPKGSKHP 1740  
Db 1681 SGEFEKRDITPTEGRSDEAGQKTSVTIPELDDNKAEGDILAEICINAMPKGSKHP 1740  
Qy 1741 FRVKKIMDVQOQASASSAPKNOLDGKKKKTPSPVKPIPQNTYTRVTRVKNADSKNNLN 1800  
Db 1741 FRVKKIMDVQOQASASSAPKNOLDGKKKKTPSPVKPIPQNTYTRVTRVKNADSKNNLN 1800  
Qy 1801 AERFVSQKSKONLKNKNDKLNNDKLVGSGFADSPHHYTPIGTPTVCFSRND 1860  
Db 1801 AERFVSQKSKONLKNKNDKLNNDKLVGSGFADSPHHYTPIGTPTVCFSRND 1860  
Qy 1861 SLSDLEDDDDVLSREKAEKAKENKSEAKVTSHTELTSNQQSANKTQAIKQFINR 1920  
Db 1861 SLSDLEDDDDVLSREKAEKAKENKSEAKVTSHTELTSNQQSANKTQAIKQFINR 1920  
Qy 1921 GQPKILOKQSTFQSSKDIPTDRAAATDEKLFPAIENTVCFSHNSLSLSDIDQENN 1980  
Db 1921 GQPKILOKQSTFQSSKDIPTDRAAATDEKLFPAIENTVCFSHNSLSLSDIDQENN 1980  
Qy 1981 NKNEPIKETEPDPSQGEPSKQASGAPKSFHVEDTPVCFSRNSLSLSIDSEDDLLQ 2040  
Db 1981 NKNEPIKETEPDPSQGEPSKQASGAPKSFHVEDTPVCFSRNSLSLSIDSEDDLLQ 2040  
Qy 2041 ECISAMPKPKKRLKDNKEXHPRNMGILGEBDLTLDKIDQRPDSEHGLSPDSENF 2100  
Db 2041 ECISAMPKPKKRLKDNKEXHPRNMGILGEBDLTLDKIDQRPDSEHGLSPDSENF 2100  
Qy 2101 WKALQEGANSIVSLHQAACLSRAQSDSDSILSKSGISLGPFFHLPDQEEKPFT 2160  
Db 2101 WKALQEGANSIVSLHQAACLSRAQSDSDSILSKSGISLGPFFHLPDQEEKPFT 2160  
Qy 2161 SNKGPRILKPEKSTLTETKIESKGIKGGKVKYKSLITGKVRNSNISEIGQWKQPLQAN 2220  
Db 2161 SNKGPRILKPEKSTLTETKIESKGIKGGKVKYKSLITGKVRNSNISEIGQWKQPLQAN 2220  
Qy 2221 MPSTSRGTMTIHIPGVNNSSTSPVSKGPPKPTPASKPSSEGTATTPRGAKPSVK 2280  
Db 2221 MPSTSRGTMTIHIPGVNNSSTSPVSKGPPKPTPASKPSSEGTATTPRGAKPSVK 2280  
Qy 2281 ELSFVARQTSIQGSSKAPSGSRDSTPSRPAQOPLSRPIQSPGRNSISPGRNGISPPN 2340  
Db 2281 ELSFVARQTSIQGSSKAPSGSRDSTPSRPAQOPLSRPIQSPGRNSISPGRNGISPPN 2340  
Qy 2341 KLSQLPRTSSPSTASTSSGSKMSYTSRQNSQOQLTKOTGLSKNASSIPRSESASKG 2400  
Db 2341 KLSQLPRTSSPSTASTSSGSKMSYTSRQNSQOQLTKOTGLSKNASSIPRSESASKG 2400

Qy 2401 LNMNMNGANKKVELSRMSTKSGSSEDRSERPVLVRQSTFIKEAPSPTLRKLBSA 2460  
Db 2401 LNMNMNGANKKVELSRMSTKSGSSEDRSERPVLVRQSTFIKEAPSPTLRKLBSA 2460  
Qy 2461 SPESLSPSRPASPTRSQAQTFVLSPSLPDMSLTHSSVQAGGWRKLPNLSPTIEYNDG 2520  
Db 2461 SPESLSPSRPASPTRSQAQTFVLSPSLPDMSLTHSSVQAGGWRKLPNLSPTIEYNDG 2520  
Qy 2521 RPAKHDHARSHSPSRSLPNNRSGTWKREHSKHSSSLPRVSTWRTGSSSILSASES 2580  
Db 2521 RPAKHDHARSHSPSRSLPNNRSGTWKREHSKHSSSLPRVSTWRTGSSSILSASES 2580  
Qy 2581 SEKAKSEDEKHVNSISGTQSKENQVSAKGWTKIKENEPSPTNSTQTVSSGATNGAES 2640  
Db 2581 SEKAKSEDEKHVNSISGTQSKENQVSAKGWTKIKENEPSPTNSTQTVSSGATNGAES 2640  
Qy 2641 KTLIYQMAPAVSKTEDVWVRIEDCPINNPRSGRPTGNTPPVIDSVSEKANPNIKOSKDN 2700  
Db 2641 KTLIYQMAPAVSKTEDVWVRIEDCPINNPRSGRPTGNTPPVIDSVSEKANPNIKOSKDN 2700  
Qy 2701 QAKQNVGNGSVPMRTVGLNRLTSFIOVDAPDQKTEIKPGQNNPVPVSETNESPIVERT 2760  
Db 2701 QAKQNVGNGSVPMRTVGLNRLTSFIOVDAPDQKTEIKPGQNNPVPVSETNESPIVERT 2760  
Qy 2761 PFSSSSSSKSHSPSGTVAARVTPFNYPNPSPRKSSADSTARSPOIPTPVNNNTKKRDKT 2820  
Db 2761 PFSSSSSSKSHSPSGTVAARVTPFNYPNPSPRKSSADSTARSPOIPTPVNNNTKKRDKT 2820  
Qy 2821 DSTESSGTSQPKHSGSYLVTSV 2843  
Db 2821 DSTESSGTSQPKHSGSYLVTSV 2843

## RESULT 4

US-08-370-235A-2  
; Sequence 2, Application US/08370235A  
; Patent No. 5910418  
; GENERAL INFORMATION:  
; APPLICANT: VOGELSTEIN, BERT  
; APPLICANT: KINZLER, KENNETH W.  
; APPLICANT: HILL, DAVID E.  
; APPLICANT: JOHNSON, KAREN A.  
; TITLE OF INVENTION: ANTIBODIES AND ASSAYS FOR DERMINING  
; TITLE OF INVENTION: MUTATIONS IN THE APC GENE  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BANNER & WITCOFF, LTD.  
; STREET: 1001 G STREET, N.W.  
; CITY: WASHINGTON  
; STATE: DC  
; COUNTRY: US  
; ZIP: 20001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/370,235A  
; FILING DATE: 01-JAN-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: KAGAN, SARAH A.  
; REGISTRATION NUMBER: 32,141  
; REFERENCE/DOCKET NUMBER: 01107.48688  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202 508 9100  
; TELEFAX: 202 508 9299  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2843 amino acids  
; TYPE: amino acid

TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-370-235A-2

Query Match 100.0%; Score 14575; DB 2; Length 2843;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2843; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MAAASYDQLLKQVEALKMENSRLRQLEDNENHLTKLETEASNKEVLKQLQGSIEDEAM	60
DB	1	MAAASYDQLLKQVEALKMENSRLRQLEDNENHLTKLETEASNKEVLKQLQGSIEDEAM	60
QY	61	ASSGOIDILRLKEINLDSNPPGVKLRKMSLSYSGREGSVSRSGECPVPMWGSFPR	120
DB	61	ASSGOIDILRLKEINLDSNPPGVKLRKMSLSYSGREGSVSRSGECPVPMWGSFPR	120
QY	121	RGFVNGSRSTGYLEFELEKERSLLIADLDKEEKDWYIAQLNLTIRIDSLPTENFSL	180
DB	121	RGFVNGSRSTGYLEFELEKERSLLIADLDKEEKDWYIAQLNLTIRIDSLPTENFSL	180
QY	181	QTDLTRQLEYEARQIRVAMEEQLGTQDMKRAQRRIARTQOIEKDLIRIQLQSQAT	240
DB	181	QTDLTRQLEYEARQIRVAMEEQLGTQDMKRAQRRIARTQOIEKDLIRIQLQSQAT	240
QY	241	EAERSSQNHETGSHDAERQEGQVGEINMATSGNGQCGSTRMDHETASVLSSSSTHSA	300
DB	241	EAERSSQNHETGSHDAERQEGQVGEINMATSGNGQCGSTRMDHETASVLSSSSTHSA	300
QY	301	PRRLTSHLGTKVEMYVYLLSMGLTHDXDMGRTLLAMSSQDSCISMROGCLPLLIQLL	360
DB	301	PRRLTSHLGTKVEMYVYLLSMGLTHDXDMGRTLLAMSSQDSCISMROGCLPLLIQLL	360
QY	361	HGNDKXSVLLGNRSCKEAPASAAJLHNIHSQPDCKRGREIRVLHLLQIRAYCETC	420
DB	361	HGNDKXSVLLGNRSCKEAPASAAJLHNIHSQPDCKRGREIRVLHLLQIRAYCETC	420
QY	421	NEWQEAHEFGMDQKNPAPVEHQICPAVCVLMKLSFDEHRRHAMNELGLQIAELQL	480
DB	421	NEWQEAHEFGMDQKNPAPVEHQICPAVCVLMKLSFDEHRRHAMNELGLQIAELQL	480
QY	481	VDCEMYGLTNDHYSITLRYAGMALTNLTFGDVANKATILCSMKGCMRALVAQLKSESDL	540
DB	481	VDCEMYGLTNDHYSITLRYAGMALTNLTFGDVANKATILCSMKGCMRALVAQLKSESDL	540
QY	541	QQVIAVLRLNSWRADVNSKTLREVGSVKALMECALEVKKESILKSVLSALWNLSAHCT	600
DB	541	QQVIAVLRLNSWRADVNSKTLREVGSVKALMECALEVKKESILKSVLSALWNLSAHCT	600
QY	601	ENKADICAVDGAFLVGLTYRSQTNLTATIESGGILIRNVSSLIATNEDHRQILRENN	660
DB	601	ENKADICAVDGAFLVGLTYRSQTNLTATIESGGILIRNVSSLIATNEDHRQILRENN	660
QY	661	CLQTLLOHLKSHSLTIVSNACGTLNLSARNPKDQBALWDMGAVSMKLNLIHSHKXIAM	720
DB	661	CLQTLLOHLKSHSLTIVSNACGTLNLSARNPKDQBALWDMGAVSMKLNLIHSHKXIAM	720
QY	721	GSAALRLNLMANRPARYKDANIMSPGSSILPSLHVHVKQKALEDAQLSETFDNIDLS	780
DB	721	GSAALRLNLMANRPARYKDANIMSPGSSILPSLHVHVKQKALEDAQLSETFDNIDLS	780
QY	781	PKASHRSKORHKQSLYGDVVDNTRHDNRSDNFTGNMTVLSPYLNLTTLVPSSSSRGS	840
DB	781	PKASHRSKORHKQSLYGDVVDNTRHDNRSDNFTGNMTVLSPYLNLTTLVPSSSSRGS	840
QY	841	LDSSRSKDRSLEREREGIGLGNVHPATENPGTSSKRGQLISTTAAQIAKWEEVSAIHTS	900
DB	841	LDSSRSKDRSLEREREGIGLGNVHPATENPGTSSKRGQLISTTAAQIAKWEEVSAIHTS	900
QY	901	QEDRSSGSTTELHCVTDERNALRRSAAHTNTYNTFKSENSNRTCSMPYAKLEYKRSS	960
DB	901	QEDRSSGSTTELHCVTDERNALRRSAAHTNTYNTFKSENSNRTCSMPYAKLEYKRSS	960
QY	961	NDLSNSVSSNDGVKGQMKPSIESYSEDDSKFCSYGQYPADLAHKIHSANHMDNDGE	1020

DB	961	NDLSNSVSSNDGVKGQMKPSIESYSEDDSKFCSYGQYPADLAHKIHSANHMDNDGE	1020
QY	1021	LDPINYSKYSDQOLNSGRQSPQSNRWARPKHIIEDBIKQSEQRQSNQSTTYPVYTE	1080
DB	1021	LDPINYSKYSDQOLNSGRQSPQSNRWARPKHIIEDBIKQSEQRQSNQSTTYPVYTE	1080
QY	1081	STDDKHLKQPHFGQGEVSPYRSGANGSETRVNSNGINQNVQSILCOEDDIEDDKP	1140
DB	1081	STDDKHLKQPHFGQGEVSPYRSGANGSETRVNSNGINQNVQSILCOEDDIEDDKP	1140
QY	1141	TNYSERYSEEOHEBERPTNYSIKYNEBKRVHDQIDYSLKYATDIPSSQKQSFSSKS	1200
DB	1141	TNYSERYSEEOHEBERPTNYSIKYNEBKRVHDQIDYSLKYATDIPSSQKQSFSSKS	1200
QY	1201	SSQSSKTEHMGSSSENTSTPSNAKRONQLHPSSAQSRSGQPOKAATCKVSSINOETIQ	1260
DB	1201	SSQSSKTEHMGSSSENTSTPSNAKRONQLHPSSAQSRSGQPOKAATCKVSSINOETIQ	1260
QY	1261	TYCVEDTPTCFSCSSLSLSAEDIEGNCQTQEADSAINTQIAIEIKKIGTRSAEDPV	1320
DB	1261	TYCVEDTPTCFSCSSLSLSAEDIEGNCQTQEADSAINTQIAIEIKKIGTRSAEDPV	1320
QY	1321	SEVPAYSQHPRTKSSRLQSSLSSESARHKAVEFPSCAKSPKSGAQTPKSPPEHYVQET	1380
DB	1321	SEVPAYSQHPRTKSSRLQSSLSSESARHKAVEFPSCAKSPKSGAQTPKSPPEHYVQET	1380
QY	1381	PLMFSRCTSVSSLDSPESRSIASSVQSEPCSGMVSGIISPSDLPDPSGQTMPSPRSKTPP	1440
DB	1381	PLMFSRCTSVSSLDSPESRSIASSVQSEPCSGMVSGIISPSDLPDPSGQTMPSPRSKTPP	1440
QY	1441	PPQTAQTKREVFNKAPTAEKRESGPKQAANAAVQRVQVLPDADTLHLFAFATESPDGF	1500
DB	1441	PPQTAQTKREVFNKAPTAEKRESGPKQAANAAVQRVQVLPDADTLHLFAFATESPDGF	1500
QY	1501	SCSSSLSALSILDPPTFOKVELRIMPVQNDNGNETESEQPKESNEQKEAEKIIDSE	1560
DB	1501	SCSSSLSALSILDPPTFOKVELRIMPVQNDNGNETESEQPKESNEQKEAEKIIDSE	1560
QY	1561	KDILLDSDDDDIIEBECIISAMPTKSSRKGKPAOTASKLPPPVARKPSQOLPVYKLLPS	1620
DB	1561	KDILLDSDDDDIIEBECIISAMPTKSSRKGKPAOTASKLPPPVARKPSQOLPVYKLLPS	1620
QY	1621	QNRLOPKHVSFTPGDDMPRVYCVVEGTPINFSTATSLSDLTIESPPNELAAGVGVRGAQ	1680
DB	1621	QNRLOPKHVSFTPGDDMPRVYCVVEGTPINFSTATSLSDLTIESPPNELAAGVGVRGAQ	1680
QY	1681	SGFEKEDTITPTEGRSTDEAOGGKTSVTIPELDDNKAEBGDILAEICINSAMPKGSHKP	1740
DB	1681	SGFEKEDTITPTEGRSTDEAOGGKTSVTIPELDDNKAEBGDILAEICINSAMPKGSHKP	1740
QY	1741	FRVKKIMDQVQOASASSAPNKNQLDGKKKKPTSPVKPIPQNTYRTRVKNADSKNNLN	1800
DB	1741	FRVKKIMDQVQOASASSAPNKNQLDGKKKKPTSPVKPIPQNTYRTRVKNADSKNNLN	1800
QY	1801	AERVFSNKKSKQNLKNSKDFNDKLPNNEDVRGSPAFDPSPHVPTPIEGTFCPSRND	1860
DB	1801	AERVFSNKKSKQNLKNSKDFNDKLPNNEDVRGSPAFDPSPHVPTPIEGTFCPSRND	1860
QY	1861	SLGSLDFDDDDVLSREKAKELRKAENKESAEKVTSHTELTSNQQSANKTQAIKQPINR	1920
DB	1861	SLGSLDFDDDDVLSREKAKELRKAENKESAEKVTSHTELTSNQQSANKTQAIKQPINR	1920
QY	1921	GQPKPILOKQSTFPQSSKDIPTDRGAATDEKLQNFALIENTPVCFSHNSLSLSLSDIOENN	1980
DB	1921	GQPKPILOKQSTFPQSSKDIPTDRGAATDEKLQNFALIENTPVCFSHNSLSLSLSDIOENN	1980
QY	1981	NKNEPIKETEPDPSQGEPSKPOASGYAPKSFVVEDTPVCFNSNSLSLSLSDIOEDDLQ	2040
DB	1981	NKNEPIKETEPDPSQGEPSKPOASGYAPKSFVVEDTPVCFNSNSLSLSLSDIOEDDLQ	2040
QY	2041	ECISSAMPKKKPSRLKGDNEKISPRNMGGILGEDJTLDKDIQRPDSEGLSPDSNFD	2100

Db 2041 ECISAMPKKKPSRLKCDNEKHSFPRNMGGLGEDLTLDLKDIOQPDSEHGLSPDSENF 2100  
 Qy 2101 WKAIQEGANSIVSSHQAACLSRQASDSDSILSKGSLGSPFHTLPDQEKPF 2160  
 Db 2101 WKAIQEGANSIVSSHQAACLSRQASDSDSILSKGSLGSPFHTLPDQEKPF 2160  
 Qy 2161 SNKGPRILKPKCKTLETETKTESSEKGIKGGKVKYKSLITGKVRNSSEISGQMKQPLQAN 2220  
 Db 2161 SNKGPRILKPKCKTLETETKTESSEKGIKGGKVKYKSLITGKVRNSSEISGQMKQPLQAN 2220  
 Qy 2221 MPSISRGRTMIHIGVRNNSSTSPVSKGPPPLKTPASKSPSEGTATTSPRCAKPSVK 2280  
 Db 2221 MPSISRGRTMIHIGVRNNSSTSPVSKGPPPLKTPASKSPSEGTATTSPRCAKPSVK 2280  
 Qy 2281 ELSPVARTSQIGSSKAPRSGRDSTPRPAQOPLSRPIQSGRNSISPGRNGISPPN 2340  
 Db 2281 ELSPVARTSQIGSSKAPRSGRDSTPRPAQOPLSRPIQSGRNSISPGRNGISPPN 2340  
 Qy 2341 KLSQLPRTSSPTASTKSSGSKMYSYTPGRQMSQQNLTKQTGLSKNASSIPRSESASKG 2400  
 Db 2341 KLSQLPRTSSPTASTKSSGSKMYSYTPGRQMSQQNLTKQTGLSKNASSIPRSESASKG 2400  
 Qy 2401 LQWANGNGANKKVELSRMSTKSSGSDRSRBPVLVROSTTIKAPSPILRKLEESA 2460  
 Db 2401 LQWANGNGANKKVELSRMSTKSSGSDRSRBPVLVROSTTIKAPSPILRKLEESA 2460  
 Qy 2461 SFESLSPSSRPASPTRSQOATPVLSPLPDMSLSTHSSVQAGGWRKLPNLSPTIEYNDG 2520  
 Db 2461 SFESLSPSSRPASPTRSQOATPVLSPLPDMSLSTHSSVQAGGWRKLPNLSPTIEYNDG 2520  
 Qy 2521 RPAKHDIARSHSPRLPINRSGTWKREHSHSSLPVSTWRTRGSSSILSASSES 2580  
 Db 2521 RPAKHDIARSHSPRLPINRSGTWKREHSHSSLPVSTWRTRGSSSILSASSES 2580  
 Qy 2581 SEKAKSDEKHVNSISGTGSKENQVSAKGTWRKIKENEPSTNSTQTVSSGATNGAES 2640  
 Db 2581 SEKAKSDEKHVNSISGTGSKENQVSAKGTWRKIKENEPSTNSTQTVSSGATNGAES 2640  
 Qy 2641 KTLIYQAPAVSKTEDVVRIEDCPINNRSGRSPGTNTPPVDSVSEKANPNIKDSKN 2700  
 Db 2641 KTLIYQAPAVSKTEDVVRIEDCPINNRSGRSPGTNTPPVDSVSEKANPNIKDSKN 2700  
 Qy 2701 QAKQNVGSPVMTVGLNRLTSFIOVDAPDQGTKEIKQNNPVVSTNETSPIVERT 2760  
 Db 2701 QAKQNVGSPVMTVGLNRLTSFIOVDAPDQGTKEIKQNNPVVSTNETSPIVERT 2760  
 Qy 2761 PFSSSSSKHSPSGTVAARVTPFNPNPSPRKSSADSTSPRPSQIPTPVNNNTKKRDSKT 2820  
 Db 2761 PFSSSSSKHSPSGTVAARVTPFNPNPSPRKSSADSTSPRPSQIPTPVNNNTKKRDSKT 2820  
 Qy 2821 DSTESSGTQPKRHSGSYLTVS 2843  
 Db 2821 DSTESSGTQPKRHSGSYLTVS 2843

RESULT 5

US-08-449-731-2

; Sequence 2, Application US/08449731

; Patent No. 6413727

; GENERAL INFORMATION:

; APPLICANT: ALBERTSEN, HANS

; ANAND, RAKESH

; CARLSON, MARY

; GRODEN, JOANNA

; HEDGE, PHILIP J.

; JOSLYN, GEOFF

; KINZLER, KENNETH

; MARKHAM, ALEXANDER F.

; NAKAMURA, YUSUKE

; THLIVERIS, ANDREW

; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC

; GENE IN COLORECTAL CANCER IN HUMANS

; NUMBER OF SEQUENCES: 102

; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Banner & Allegretti, LTD  
 ; STREET: 1001 G Street, NW  
 ; CITY: Washington  
 ; STATE: D.C.  
 ; COUNTRY: USA  
 ; ZIP: 20001-4598  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/449,731  
 ; FILING DATE: 25-May-1995  
 ; CLASSIFICATION: <Unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/289,548  
 ; FILING DATE: 12-AUG-1994  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Kagan, Sarah A.  
 ; REGISTRATION NUMBER: 32,141  
 ; REFERENCE/DOCKET NUMBER: 1107.46943  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 202-508-9100  
 ; TELEFAX: 202-508-9299  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 2843 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
 ; US-08-449-731-2

Query Match 100.0%; Score 14575; DB 4; Length 2843;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2843; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAAASYDQLLKQVAKLVKNSLNRLQEDNSNHLTKLETSANMKVYLKQLQGSIEDEAM 60  
 Db 1 MAAASYDQLLKQVAKLVKNSLNRLQEDNSNHLTKLETSANMKVYLKQLQGSIEDEAM 60  
 Qy 61 ASSGOIDLLELRLKEINLDSSNFPVGLRSKMSLSYSGRSVSSRSRSGECSVPVMSGFPR 120  
 Db 61 ASSGOIDLLELRLKEINLDSSNFPVGLRSKMSLSYSGRSVSSRSRSGECSVPVMSGFPR 120  
 Qy 121 RGFVNGSRESSTGYLEELKERSILLADLDKEEKDWMYVLAQLNLTKRIIDSLPTENFSL 180  
 Db 121 RGFVNGSRESSTGYLEELKERSILLADLDKEEKDWMYVLAQLNLTKRIIDSLPTENFSL 180  
 Qy 181 QTDLTRRQLEYEARQIRVAMEEQLGTCDMEKRAQRRIARIQIEKDILRIQLLSQAT 240  
 Db 181 QTDLTRRQLEYEARQIRVAMEEQLGTCDMEKRAQRRIARIQIEKDILRIQLLSQAT 240  
 Qy 241 EAERSSQKHETGSHDAERQNEGQVGEINMATSNGQGGSTTRMDHETASVLSSTHSA 300  
 Db 241 EAERSSQKHETGSHDAERQNEGQVGEINMATSNGQGGSTTRMDHETASVLSSTHSA 300  
 Qy 301 PRRLTSHLGTKEVMVYSLLSMLGTHDKDMSRTLLAMSSQDSCISMRQSGCPLLIQLL 360  
 Db 301 PRRLTSHLGTKEVMVYSLLSMLGTHDKDMSRTLLAMSSQDSCISMRQSGCPLLIQLL 360  
 Qy 361 HGNDKDSVLLGNRSGSKERARASAAAHNIHSDPKRGERETRVHLLEQIYACETC 420  
 Db 361 HGNDKDSVLLGNRSGSKERARASAAAHNIHSDPKRGERETRVHLLEQIYACETC 420  
 Qy 421 WEWQEAHEPGMDQDKNPMPAPVEHQICPAVCVLMKLSFDEEHRHAMNELGGLQAIALLQ 480  
 Db 421 WEWQEAHEPGMDQDKNPMPAPVEHQICPAVCVLMKLSFDEEHRHAMNELGGLQAIALLQ 480  
 Qy 481 VDCMEGLTNDHYSITLRRYAGVALLTLTFGDVANKATLCSMKGCWRAVLAQLKSEEDL 540

481 VDCMYGLTNDHYSITLRRYAGMALTNLTFGDVANKATLCSMKOCMALVAQLKSESED 540  
481 QOVIASVLRLNLSWRADVNSKKTLEVGSVKALMECALEVKKESTLKSVLKALWNLSAHCT 600  
541 QOVIASVLRLNLSWRADVNSKKTLEVGSVKALMECALEVKKESTLKSVLKALWNLSAHCT 600  
601 ENKADICAVDGCALAFVGLTLYRSOTNTLAIIESGGGILRNVSLSLIATNEDHROILRENN 660  
601 ENKADICAVDGCALAFVGLTLYRSOTNTLAIIESGGGILRNVSLSLIATNEDHROILRENN 660  
661 CLQTLQLHLKSHSLTIVSNAGCTLWNLARNPKQOEALWDMGVSMKLNLIHSKHKMIAM 720  
661 CLQTLQLHLKSHSLTIVSNAGCTLWNLARNPKQOEALWDMGVSMKLNLIHSKHKMIAM 720  
721 GSAALRLNLMANRPKAKYKADANIMSPGSSLSLHVYRKQKALEAELDAOHLSETFNDNLNLS 780  
721 GSAALRLNLMANRPKAKYKADANIMSPGSSLSLHVYRKQKALEAELDAOHLSETFNDNLNLS 780  
781 PKASHRSKORHKQSLGYDVTNHRDNRDNFNTGNMTNMTVLSPLYNTTIVLPSSSSSRGS 840  
781 PKASHRSKORHKQSLGYDVTNHRDNRDNFNTGNMTNMTVLSPLYNTTIVLPSSSSSRGS 840  
841 LDSSRSKDRSLERBRGIGLGNYPHATENPGTSSKRGLOISTTAAQIAKYMEEVSALHTS 900  
841 LDSSRSKDRSLERBRGIGLGNYPHATENPGTSSKRGLOISTTAAQIAKYMEEVSALHTS 900  
901 QEDRSSGSTTELHCVTDERNALRRSSAAHTSNNTYNTFKSENSNRTCSMPYAKLEYKRSS 960  
901 QEDRSSGSTTELHCVTDERNALRRSSAAHTSNNTYNTFKSENSNRTCSMPYAKLEYKRSS 960  
961 NDSLNSVSNDDGYKRGQMKPSIESYEDDESKFCYQYQPADLAHKIHSANHMDNDGE 1020  
961 NDSLNSVSNDDGYKRGQMKPSIESYEDDESKFCYQYQPADLAHKIHSANHMDNDGE 1020  
1021 LDTPNYSIKYSDQELNSGRQSPSONERKAPKHIIEDEIKQSEORQSRNOSTTYPVYTE 1080  
1021 LDTPNYSIKYSDQELNSGRQSPSONERKAPKHIIEDEIKQSEORQSRNOSTTYPVYTE 1080  
1081 STDDXHLAFQPHFGQOECVSPYRSGANGSETRNRYGSHNGINQNVQSLSQCEDDEDYDDKP 1140  
1081 STDDXHLAFQPHFGQOECVSPYRSGANGSETRNRYGSHNGINQNVQSLSQCEDDEDYDDKP 1140  
1141 TNYSRYSBEEHQBEEERPTNYSIKYNEEKRHVDQIDYSLKYATDIPSSOKQSFSPKS 1200  
1141 TNYSRYSBEEHQBEEERPTNYSIKYNEEKRHVDQIDYSLKYATDIPSSOKQSFSPKS 1200  
1201 SSGQSKTEHMSSENSTPSSNAKRONLHPSSAQSRSGOPOKAATCKVSSINQETIQ 1260  
1201 SSGQSKTEHMSSENSTPSSNAKRONLHPSSAQSRSGOPOKAATCKVSSINQETIQ 1260  
1261 TYCVEDTPICFGRCSLSLSAEDIEGQNTQTEADSAANTLQIAEIKGKIGTRSAEDPV 1320  
1261 TYCVEDTPICFGRCSLSLSAEDIEGQNTQTEADSAANTLQIAEIKGKIGTRSAEDPV 1320  
1321 SEVPVAVSOHPRTKSLFLOGSSLSSESAHKAVERFPSSGAKSPSKGAOTPKSPPEHYVOET 1380  
1321 SEVPVAVSOHPRTKSLFLOGSSLSSESAHKAVERFPSSGAKSPSKGAOTPKSPPEHYVOET 1380  
1381 PLMFSRCTSVSLSDFPESRSIASSVQSEPCSGWVGIIISPSLDPSPGQTPPPSRSKTTP 1440  
1381 PLMFSRCTSVSLSDFPESRSIASSVQSEPCSGWVGIIISPSLDPSPGQTPPPSRSKTTP 1440  
1441 PPPQTAQTRVPKPKAPTAERESGPKQAANVAORVQVLPDADTLHLHATESTPDGF 1500  
1441 PPPQTAQTRVPKPKAPTAERESGPKQAANVAORVQVLPDADTLHLHATESTPDGF 1500  
1501 SCSSLSALSALDEPFFOKOVELRIMPVQENDNGNETESEQPKSENEQKEAEXTIDSE 1560  
1501 SCSSLSALSALDEPFFOKOVELRIMPVQENDNGNETESEQPKSENEQKEAEXTIDSE 1560  
1561 KOLLDDSDDDDEIILBECIIISAMPTKSRKGGKPKAQTASKLPPPVARKPSQLPVYKLLPS 1620  
1561 KOLLDDSDDDDEIILBECIIISAMPTKSRKGGKPKAQTASKLPPPVARKPSQLPVYKLLPS 1620

1621 QNELQOKHVSTPFGDDMPRVYCVBGTPIINFSTATSLSDLTIESPPNELAAGEVGGQA 1680  
1621 QNELQOKHVSTPFGDDMPRVYCVBGTPIINFSTATSLSDLTIESPPNELAAGEVGGQA 1680  
1681 SGBFEKREDTIPTEGRSTDEAOGKTSSTVPIELDDNKAEEGDILAEICINSAMPKGSHP 1740  
1681 SGBFEKREDTIPTEGRSTDEAOGKTSSTVPIELDDNKAEEGDILAEICINSAMPKGSHP 1740  
1741 FRVKYIMQVQQAASASSAPNKNQLDGKKKKTSPVKPIPONTETRYTRVRKNAADSKNLN 1800  
1741 FRVKYIMQVQQAASASSAPNKNQLDGKKKKTSPVKPIPONTETRYTRVRKNAADSKNLN 1800  
1801 AERFVSDNKKDKONLKNNSKDFNDKLPNNEDVRGSPAFDPHHPHTPIEGTPYCFERN 1860  
1801 AERFVSDNKKDKONLKNNSKDFNDKLPNNEDVRGSPAFDPHHPHTPIEGTPYCFERN 1860  
1861 SLSLDPDDDDVLSREKAEELRAKAKENSEAKVTSHTELTNSQOQANKTOIAKQPINR 1920  
1861 SLSLDPDDDDVLSREKAEELRAKAKENSEAKVTSHTELTNSQOQANKTOIAKQPINR 1920  
1921 GQPKPILQKOSTPQSSKDIIPDRGAATDEKLOFNFAIENTPVCFSHNSLSLSDIDQENN 1980  
1921 GQPKPILQKOSTPQSSKDIIPDRGAATDEKLOFNFAIENTPVCFSHNSLSLSDIDQENN 1980  
1981 NKENEPIKETETPPDSQCFSPSKQASGYAPKSFHVEDTPVCFGRNSLSLSIDSEDDLQ 2040  
1981 NKENEPIKETETPPDSQCFSPSKQASGYAPKSFHVEDTPVCFGRNSLSLSIDSEDDLQ 2040  
2041 ECISAMPKPKSRILKGDNEKHSPRNMGILGEDTLDLKOIQRDPDSEHGLSPDSENF 2100  
2041 ECISAMPKPKSRILKGDNEKHSPRNMGILGEDTLDLKOIQRDPDSEHGLSPDSENF 2100  
2101 WKAIQEGANSIVSLSHQAACCLSRQASDSDSILSLXSGISLGSPPFHLTPQBEKPP 2160  
2101 WKAIQEGANSIVSLSHQAACCLSRQASDSDSILSLXSGISLGSPPFHLTPQBEKPP 2160  
2161 SNKGPRILKPKGKSTLTETKIBESKIGKGGKVKYKSLITGKVRNSSEISGQWKQLOAN 2220  
2161 SNKGPRILKPKGKSTLTETKIBESKIGKGGKVKYKSLITGKVRNSSEISGQWKQLOAN 2220  
2221 MPSTSRGRTMIHPIGVNNSSTSPVSKGPPKLTTPASKSPSEGOATATSPRGAKEPVKS 2280  
2221 MPSTSRGRTMIHPIGVNNSSTSPVSKGPPKLTTPASKSPSEGOATATSPRGAKEPVKS 2280  
2281 ELSFVARQTSQIGGSKKAPSRGSRDSTPSRPAQQLSRPIQSPGRNSISPGRNGISPPN 2340  
2281 ELSFVARQTSQIGGSKKAPSRGSRDSTPSRPAQQLSRPIQSPGRNSISPGRNGISPPN 2340  
2341 KLSQLPRTSPSTASTKSSGSGKMSYTSPCROWSQOQLTKOTGLSKNASSIPSESEASKG 2400  
2341 KLSQLPRTSPSTASTKSSGSGKMSYTSPCROWSQOQLTKOTGLSKNASSIPSESEASKG 2400  
2401 LNQMNGNGANKKVELSRMSTKSSGSESDRSPVLRQSTFIKEAPSTLRRKLEESA 2460  
2401 LNQMNGNGANKKVELSRMSTKSSGSESDRSPVLRQSTFIKEAPSTLRRKLEESA 2460  
2461 SFESLSPSPSPAPSTRSOAQTIVLSPLPMSLSTHSSVOAGGWRKLPNLSPTIENDG 2520  
2461 SFESLSPSPSPAPSTRSOAQTIVLSPLPMSLSTHSSVOAGGWRKLPNLSPTIENDG 2520  
2521 RPAKRDHIAHSHSPRLPINRSGTWKREHSHSSSLPRVSTWRTTGSSSSILSASSES 2580  
2521 RPAKRDHIAHSHSPRLPINRSGTWKREHSHSSSLPRVSTWRTTGSSSSILSASSES 2580  
2581 SEKAKSEDEKHVNSISGTKQSKENQVSAKWTWKIKENEFSPNTSQTIVSSGATNGAES 2640  
2581 SEKAKSEDEKHVNSISGTKQSKENQVSAKWTWKIKENEFSPNTSQTIVSSGATNGAES 2640  
2641 KTLIYQAPAVSKTDEWVRIEDCPINNPRSGSPGTCNTPPVIDSVSEKANPNIKDSKN 2700  
2641 KTLIYQAPAVSKTDEWVRIEDCPINNPRSGSPGTCNTPPVIDSVSEKANPNIKDSKN 2700

QY 2701 QAKQVNGSVPMRTVGLNRLTSFIQVDAPDQKGTBTKPGQNNPVPVSETNBSPIVERT 2760  
 Db 2701 QAKQVNGSVPMRTVGLNRLTSFIQVDAPDQKGTBTKPGQNNPVPVSETNBSPIVERT 2760  
 QY 2761 PFSSSSSKSHSPSGTVAARVTPPNYNPSPKSSADSTARSPOIPTPVNNNTKKRDSKT 2820  
 Db 2761 PFSSSSSKSHSPSGTVAARVTPPNYNPSPKSSADSTARSPOIPTPVNNNTKKRDSKT 2820  
 QY 2821 DSTESSGTQSPKRHSGLVTSV 2843  
 Db 2821 DSTESSGTQSPKRHSGLVTSV 2843  
 RESULT 6  
 US-08-452-655B-2  
 ; Sequence 2, Application US/08452655B  
 ; Patent No. 5783666  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ALBERTSEN, HANS  
 ; APPLICANT: ANAND, RAKESH  
 ; APPLICANT: CARLSON, MARY  
 ; APPLICANT: GRODEN, JOANNA  
 ; APPLICANT: HEDGE, PHILIP J.  
 ; APPLICANT: JOSLYN, GEOFF  
 ; APPLICANT: KINZLER, KENNETH  
 ; APPLICANT: MARKHAM, ALEXANDER F.  
 ; APPLICANT: NAKAMURA, YUSUKE  
 ; APPLICANT: THLIVERIS, ANDREW  
 ; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC  
 ; TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS  
 ; NUMBER OF SEQUENCES: 102  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Banner & Witcoff, Ltd.  
 ; STREET: 1001 G Street, NW  
 ; CITY: Washington  
 ; STATE: D.C.  
 ; COUNTRY: USA  
 ; ZIP: 20001-4598  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/452,655B  
 ; FILING DATE: 25-MAY-1995  
 ; CLASSIFICATION: 530  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/289,548  
 ; FILING DATE: 12-AUG-1994  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/741,940  
 ; FILING DATE: 08-AUG-1991  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Kagen, Sarah A.  
 ; REGISTRATION NUMBER: 32,141  
 ; REFERENCE/DOCKET NUMBER: 1107.49964  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 202-508-9100  
 ; TELEFAX: 202-508-9299  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 2843 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-452-655B-2  
 Query Match 99.7%; Score 14533; DB 1; Length 2843;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 2836; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 MAAASYDQLLKQVEALKMENSNLRLQTEASNMKEVLKQLOGSIEDEAM 60

Db 1 MAAASYDQLLKQVEALKMENSNLRLQTEASNMKEVLKQLOGSIEDEAM 60  
 QY 61 ASSGOIDLLERLKEINLNDSSNFPVGVKLRSMKLSYSGRSGSVSRSGSCECPVPMWGSPPR 120  
 Db 61 ASSGOIDLLERLKEINLNDSSNFPVGVKLRSMKLSYSGRSGSVSRSGSCECPVPMWGSPPR 120  
 QY 121 RGFVNGSRESTGYLBELEKERSLLADLDKEBEKDWYQAQLNLTNRIDSLPTENPFL 180  
 Db 121 RGFVNGSRESTGYLBELEKERSLLADLDKEBEKDWYQAQLNLTNRIDSLPTENPFL 180  
 QY 181 QDILTRROLEYEARQIRVAMEEQIGTODMEKRAQRRIARIQQIEKOLLIRQLQSOAT 240  
 Db 181 QDILTRROLEYEARQIRVAMEEQIGTODMEKRAQRRIARIQQIEKOLLIRQLQSOAT 240  
 QY 241 EAERSSQNKHETGSHDAERQNEGGQVGEINMATSGNGQGSTTRMDHETASVLSSTHSA 300  
 Db 241 EAERSSQNKHETGSHDAERQNEGGQVGEINMATSGNGQGSTTRMDHETASVLSSTHSA 300  
 QY 301 PRLTSHLGTKEVYVYLLSMLGTHDKDMSRTLLAMSSQSDSCISMRQSCPLILLQLL 360  
 Db 301 PRLTSHLGTKEVYVYLLSMLGTHDKDMSRTLLAMSSQSDSCISMRQSCPLILLQLL 360  
 QY 361 HGNDKDSVLLGNRSGSKEARARASAAALHNIHSDPDDKRGRRREIRVLHLEQIRAYCETC 420  
 Db 361 HGNDKDSVLLGNRSGSKEARARASAAALHNIHSDPDDKRGRRREIRVLHLEQIRAYCETC 420  
 QY 421 WEQBAHEPGMDQKNPMPAPVEHQICPAVCVLMKLSFDEEHRHAMNELGQLIAELLQ 480  
 Db 421 WEQBAHEPGMDQKNPMPAPVEHQICPAVCVLMKLSFDEEHRHAMNELGQLIAELLQ 480  
 QY 481 VDCMYGLTNDHYSITLRRYAGMALTNLTGFDVANKATLCSMKCMRALVAQLKSEEDL 540  
 Db 481 VDCMYGLTNDHYSITLRRYAGMALTNLTGFDVANKATLCSMKCMRALVAQLKSEEDL 540  
 QY 541 QQVIASVLRNLSWRADVNSKTLREVSGVKALMECALEVKESTLKSVALMNLASHCT 600  
 Db 541 QQVIASVLRNLSWRADVNSKTLREVSGVKALMECALEVKESTLKSVALMNLASHCT 600  
 QY 601 ENKADICAVDICALAPLVGTLTYRSQNTNTLAIIESGGGILNRVSSLIATNEDHROIENN 660  
 Db 601 ENKADICAVDICALAPLVGTLTYRSQNTNTLAIIESGGGILNRVSSLIATNEDHROIENN 660  
 QY 661 CLQTLQHLKSHSLTIIVSNACGTLNLSARNPKDQEAALWDWGVSMKLNLIHSHKMIAM 720  
 Db 661 CLQTLQHLKSHSLTIIVSNACGTLNLSARNPKDQEAALWDWGVSMKLNLIHSHKMIAM 720  
 QY 721 GSAALNLMANRPAKYKDANINSPGSSLPDLHVKKQKALEAELDAQHLSETPDNIOLS 780  
 Db 721 GSAALNLMANRPAKYKDANINSPGSSLPDLHVKKQKALEAELDAQHLSETPDNIOLS 780  
 QY 781 PKASHRSKQRHKQSLGYDYVFDNTRHDDNRSDNFTNMTVLSPLYNTTLPSSSSSRGS 840  
 Db 781 PKASHRSKQRHKQSLGYDYVFDNTRHDDNRSDNFTNMTVLSPLYNTTLPSSSSSRGS 840  
 QY 841 LDSSRSKDRSLERERIGIGNVHPATENFGTSKRGLOISTTAAQIAKWEVSAIHTS 900  
 Db 841 LDSSRSKDRSLERERIGIGNVHPATENFGTSKRGLOISTTAAQIAKWEVSAIHTS 900  
 QY 901 QEDRSSGSTTTELHCVTDERNALRRSSAAHSTHNTYNTFKSENSNRCTCMFYAKLEYKRSS 960  
 Db 901 QEDRSSGSTTTELHCVTDERNALRRSSAAHSTHNTYNTFKSENSNRCTCMFYAKLEYKRSS 960  
 QY 961 NDLSNVSSNDGYGKRCQMKPSIESYSEDEDESKFCYGYPADLAHKAHSHANMDNDGE 1020  
 Db 961 NDLSNVSSNDGYGKRCQMKPSIESYSEDEDESKFCYGYPADLAHKAHSHANMDNDGE 1020  
 QY 1021 LDTPINYSKYSDEQLNSGRQSPSONERWARPKHIIDEIKQSEQRQSRNQSTTYPYTE 1080  
 Db 1021 LDTPINYSKYSDEQLNSGRQSPSONERWARPKHIIDEIKQSEQRQSRNQSTTYPYTE 1080  
 QY 1081 STDDKHLKPFHQFCQSCVSPYRSRGANGSETNRVSGNHGINQVNSQLCQEDDDEDDKP 1140



Db 1081 STDDKHLKQPHFGQOCVSPVRSRGANGSETNRVGSNHHGINQNVQSCLQEDDDYEDDKP 1140  
QY 1141 TNYSEYSEEEHEEERPTNYSIIKYNEEKRHVDQPIDYSLKYATDIPSSQKQSFSEFSKS 1200  
Db 1141 TNYSEYSEEEHEEERPTNYSIIKYNEEKRHVDQPIDYSLKYATDIPSSQKQSFSEFSKS 1200  
QY 1201 SGOSSKTEHMSSENSTSPSNKRONQOLHPSQAQSGOPOKAATCKYSSINQETIQ 1260  
Db 1201 SGOSSKTEHMSSENSTSPSNKRONQOLHPSQAQSGOPOKAATCKYSSINQETIQ 1260  
QY 1261 TYCVEDTPICFGRCSLSLSLAEDEIGCNQTOQADGANTLQIAEIKGKIGTRSAEDPV 1320  
Db 1261 TYCVEDTPICFGRCSLSLSLAEDEIGCNQTOQADGANTLQIAEIKGKIGTRSAEDPV 1320  
QY 1321 SEVPASVQHPRTKSRRLQSSLSSESARHKAVEFPSSGAKSPKSGAQTPKSPPEHYVOET 1380  
Db 1321 SEVPASVQHPRTKSRRLQSSLSSESARHKAVEFPSSGAKSPKSGAQTPKSPPEHYVOET 1380  
QY 1381 PLMFGRCTSVSLDSDFESRSIASSVQSEPCSGWVGIIISPSDLPSQCTMPPPSRSKTPP 1440  
Db 1381 PLMFGRCTSVSLDSDFESRSIASSVQSEPCSGWVGIIISPSDLPSQCTMPPPSRSKTPP 1440  
QY 1441 PPQTAQTKREVPKNAKAPTAEKREGPKOAAVNAAVQVQLPDAADTLHLHFAESTPDGPF 1500  
Db 1441 PPQTAQTKREVPKNAKAPTAEKREGPKOAAVNAAVQVQLPDAADTLHLHFAESTPDGPF 1500  
QY 1501 SCSSLSLSLDBPFIQKDVLRIMPVQOENDNGNETESEQPKESNENQKEAEKTIIDSE 1560  
Db 1501 SCSSLSLSLDBPFIQKDVLRIMPVQOENDNGNETESEQPKESNENQKEAEKTIIDSE 1560  
QY 1561 KOLLDDSDDDDEIIELECIISAMPTKSSRKGGKPAQTASKLPPPVARPKPSQLPVYKLLPS 1620  
Db 1561 KOLLDDSDDDDEIIELECIISAMPTKSSRKGGKPAQTASKLPPPVARPKPSQLPVYKLLPS 1620  
QY 1621 QNRLOQKXVSTPGDDMPRVYCVGTPINFSTATSLDLTIESPNELAAGEVGRGAQ 1680  
Db 1621 QNRLOQKXVSTPGDDMPRVYCVGTPINFSTATSLDLTIESPNELAAGEVGRGAQ 1680  
QY 1681 SEFEKEDTIPTEGRSTDBAQGGKTSVTPILPDNKAEBEGDILAEICINSAMPKCKSHK 1740  
Db 1681 SEFEKEDTIPTEGRSTDBAQGGKTSVTPILPDNKAEBEGDILAEICINSAMPKCKSHK 1740  
QY 1741 FRVKIMDOVQOASASSAPNKNQLDGKKKTSVPKPIQNTYRTRVRKNKADSKNNLN 1800  
Db 1741 FRVKIMDOVQOASASSAPNKNQLDGKKKTSVPKPIQNTYRTRVRKNKADSKNNLN 1800  
QY 1801 AERVFSDNKKQKQNLKNNKDFNDKLPNNEDRVGSGFAFDSPHYTPTEGTPCYFSRND 1860  
Db 1801 AERVFSDNKKQKQNLKNNKDFNDKLPNNEDRVGSGFAFDSPHYTPTEGTPCYFSRND 1860  
QY 1861 SLSSLDFFDDDDVDLREKAEALREKAKENKESAEKVTSHTELTSNQOSANKTQAIKQPINR 1920  
Db 1861 SLSSLDFFDDDDVDLREKAEALREKAKENKESAEKVTSHTELTSNQOSANKTQAIKQPINR 1920  
QY 1921 GQPKPILOKSTFPOSSKDIIPRGATDEKLOFAIENTPVCFSHNSLSLSLSDIDQENN 1980  
Db 1921 GQPKPILOKSTFPOSSKDIIPRGATDEKLOFAIENTPVCFSHNSLSLSLSDIDQENN 1980  
QY 1981 NKENEPIKETEPPDSOGEPKQASGYAPKSPHVEDTPVCFGRNSLSLSLSDSDDDLQ 2040  
Db 1981 NKENEPIKETEPPDSOGEPKQASGYAPKSPHVEDTPVCFGRNSLSLSLSDSDDDLQ 2040  
QY 2041 ECISAMPKKKPSRLJGKNEXHSPNMGILGEDLTLDKIQPDSHGLSPDSENF 2100  
Db 2041 ECISAMPKKKPSRLJGKNEXHSPNMGILGEDLTLDKIQPDSHGLSPDSENF 2100  
QY 2101 WKAIQEGANSIVSSLHQAAAACLSRQASDSDSILSLKSGISLSPHILTPDQEKFT 2160  
Db 2101 WKAIQEGANSIVSSLHQAAAACLSRQASDSDSILSLKSGISLSPHILTPDQEKFT 2160  
QY 2161 SNKGPRILKPGKSTLETKKIESSEKIGKGGKVKYKSLITGKVRNSEISGOMKQLOAN 2220  
Db 2161 SNKGPRILKPGKSTLETKKIESSEKIGKGGKVKYKSLITGKVRNSEISGOMKQLOAN 2220

QY 2221 MFSISRGRWTMIHPIGVNRNSSSTSPVSKKGPPLKTPAKSKSPSEGTATTSPPRAKPSVKS 2280  
Db 2221 MFSISRGRWTMIHPIGVNRNSSSTSPVSKKGPPLKTPAKSKSPSEGTATTSPPRAKPSVKS 2280  
QY 2281 ELSVPARQTSOIGGSSKAPSRSGRSTPSRPAQPLSRPIQSPGRNSISPGRNGISPPN 2340  
Db 2281 ELSVPARQTSOIGGSSKAPSRSGRSTPSRPAQPLSRPIQSPGRNSISPGRNGISPPN 2340  
QY 2341 KLSQLPTTSPSTASTKSSGSGKMSYTPGRQWSQQNLTKQTGLSKNASSIPRSEASKG 2400  
Db 2341 KLSQLPTTSPSTASTKSSGSGKMSYTPGRQWSQQNLTKQTGLSKNASSIPRSEASKG 2400  
QY 2401 LQNMNGNGANKKVELSRMSSTKSSSESDESRPVLVROSTFIKEAPSTLRRKLEESA 2460  
Db 2401 LQNMNGNGANKKVELSRMSSTKSSSESDESRPVLVROSTFIKEAPSTLRRKLEESA 2460  
QY 2461 SFESLSPSSPASPTRSQATPVLSPSLPDMWSLSTHSSVQAGGWRKLPPLSPTIEYNDG 2520  
Db 2461 SFESLSPSSPASPTRSQATPVLSPSLPDMWSLSTHSSVQAGGWRKLPPLSPTIEYNDG 2520  
QY 2521 RPAKGRHDIARSHGESPSRLPINRSGTWKREHSHSSLPREVSTWRTTSSSSILSASSES 2580  
Db 2521 RPAKGRHDIARSHGESPSRLPINRSGTWKREHSHSSLPREVSTWRTTSSSSILSASSES 2580  
QY 2581 SEKAKSEDEKHVNSISGTQKSKENOVSAKGTWKIKENEPSTNSTQTVSSGATNGAES 2640  
Db 2581 SEKAKSEDEKHVNSISGTQKSKENOVSAKGTWKIKENEPSTNSTQTVSSGATNGAES 2640  
QY 2641 KTLIYQMAPAVSKTDEWVRIEDPCPINNPRSGSPGTNTPTPVIDSVSEKANPNIKDSKDN 2700  
Db 2641 KTLIYQMAPAVSKTDEWVRIEDPCPINNPRSGSPGTNTPTPVIDSVSEKANPNIKDSKDN 2700  
QY 2701 QAKQNVGSGVPMRTVGLNRLTSFOVDAPDQKGEIKPGQNNPVVSTNESPIVER 2760  
Db 2701 QAKQNVGSGVPMRTVGLNRLTSFOVDAPDQKGEIKPGQNNPVVSTNESPIVER 2760  
QY 2761 PFSSSSSSKHSPPSGTVAARVTPFNPNPRKSSADTSARPQIPTFVNNTKKRDSKT 2820  
Db 2761 PFSSSSSSKHSPPSGTVAARVTPFNPNPRKSSADTSARPQIPTFVNNTKKRDSKT 2820  
QY 2821 DSTESSGTQSPKRHSGSYLTVSV 2843  
Db 2821 DSTESSGTQSPKRHSGSYLTVSV 2843

RESULT 7  
US-08-452-655B-7  
; Sequence 7, Application US/08452655B  
; Patent No. 5783666  
; GENERAL INFORMATION:  
; APPLICANT: ALBERTSEN, HANS  
; APPLICANT: ANAND, RAKESH  
; APPLICANT: CARLSON, MARY  
; APPLICANT: GRODEN, JOANNA  
; APPLICANT: HEDGE, PHILIP J.  
; APPLICANT: JOSLYN, GEOFF  
; APPLICANT: KINZLER, KENNETH  
; APPLICANT: MARKHAM, ALEXANDER F.  
; APPLICANT: NAKAMURA, YUSUKE  
; APPLICANT: THLIVERIS, ANDREW  
; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC  
; TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS  
; NUMBER OF SEQUENCES: 102  
; CORRESPONDENCE ADDRESS:  
; STREET: 1001 G Street, NW  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20001-4598  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/452,655B  
 ; FILING DATE: 25-MAY-1995  
 ; CLASSIFICATION: 530  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/289,548  
 ; FILING DATE: 12-AUG-1994  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/741,940  
 ; FILING DATE: 08-AUG-1991  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Kagan, Sarah A. 32,141  
 ; REGISTRATION NUMBER: 32,141  
 ; REFERENCE/DOCKET NUMBER: 1107.49964  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 202-508-9100  
 ; TELEFAX: 202-508-9299  
 ; INFORMATION FOR SEQ ID NO: 7:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 2843 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; HYPOTHEetical: YES  
 ; ANTI-SENSE: NO  
 ; US-08-452-655B-7

Query Match 99.7%; Score 14533; DB 1; Length 2843;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 2836; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY	1	MAAASYDQLLKQVEALWENSLRQELSDNSHLTKLETSASNKEVYLKQCSIDEAM	60
DB	1	MAAASYDQLLKQVEALWENSLRQELSDNSHLTKLETSASNKEVYLKQCSIDEAM	60
QY	61	ASSGQIDLLERLKEINLDSNFPQVGLRSKMSLSYSGRSGSVSSRSGECPVPMGSPFR	120
DB	61	ASSGQIDLLERLKEINLDSNFPQVGLRSKMSLSYSGRSGSVSSRSGECPVPMGSPFR	120
QY	121	RGFVNGSRSTGYLEELKERSLLADLDKEEKDWYAAQLNLTFRIDSLPLTENFSL	180
DB	121	RGFVNGSRSTGYLEELKERSLLADLDKEEKDWYAAQLNLTFRIDSLPLTENFSL	180
QY	181	QTDLTRQLEEARQIRVAMEEQLGTQDMKEKRAORRIARIQIEKDLIRLQLOSQAT	240
DB	181	QTDLTRQLEEARQIRVAMEEQLGTQDMKEKRAORRIARIQIEKDLIRLQLOSQAT	240
QY	241	EAERSSQNKHETGSHDAERNEGQGVGEINATSGNGCGSTTRMDHETASVLSGSSSTHSA	300
DB	241	EAERSSQNKHETGSHDAERNEGQGVGEINATSGNGCGSTTRMDHETASVLSGSSSTHSA	300
QY	301	PRRLTSHLGTKEVMYSLMLGTHDKDDMSRTLAMSSQDSCISNRQSGCPLLIQLL	360
DB	301	PRRLTSHLGTKEVMYSLMLGTHDKDDMSRTLAMSSQDSCISNRQSGCPLLIQLL	360
QY	361	HCNDKDSVLLGNRSGSKEARASAAALHNTIHGQPDQKRRIRVHLLEQIRAYCETC	420
DB	361	HCNDKDSVLLGNRSGSKEARASAAALHNTIHGQPDQKRRIRVHLLEQIRAYCETC	420
QY	421	WEWQEAHPGMDQKNPMPAPVEHQICPAVCVLMKLSFDEHRRHANNELGGLQAIALLQ	480
DB	421	WEWQEAHPGMDQKNPMPAPVEHQICPAVCVLMKLSFDEHRRHANNELGGLQAIALLQ	480
QY	481	VDCEMYGLTNDHYITLRRVAGMALTNLTGFDVANKATLCSMKGCMFALVAQLKSESED	540
DB	481	VDCEMYGLTNDHYITLRRVAGMALTNLTGFDVANKATLCSMKGCMFALVAQLKSESED	540
QY	541	QQVIVASVLRNLSWRADVNSKKTLEREVGSVKALMECALEVKKESTLKSIVLSALWLSAHCT	600
DB	541	QQVIVASVLRNLSWRADVNSKKTLEREVGSVKALMECALEVKKESTLKSIVLSALWLSAHCT	600
QY	601	ENKADICAVDGLALAPLVGTLVRSQNTNLAIIESGGGILRVNSSLIATNEHROILRENN	660
DB	601	ENKADICAVDGLALAPLVGTLVRSQNTNLAIIESGGGILRVNSSLIATNEHROILRENN	660
QY	661	CLQTLQLKLSHSLTIVSNAGCTLWLSARNPQKQALWMDMGAVSMKLNLIHSHKHMIAM	720
DB	661	CLQTLQLKLSHSLTIVSNAGCTLWLSARNPQKQALWMDMGAVSMKLNLIHSHKHMIAM	720
QY	721	GSAAALRNLMANRPKAKYKDANIMPGSSLPSSLHVRKQKALBAELDAQHLSFTFNIDLS	780
DB	721	GSAAALRNLMANRPKAKYKDANIMPGSSLPSSLHVRKQKALBAELDAQHLSFTFNIDLS	780
QY	781	PKASHRSKQRHKQSLYGDYVDFTRHDDNRSDNFTNMTVLSPYLNTTVLPSSSSSRGS	840
DB	781	PKASHRSKQRHKQSLYGDYVDFTRHDDNRSDNFTNMTVLSPYLNTTVLPSSSSSRGS	840
QY	841	LDSSRSSEKDRSLERERIGLGNYPHATENPCTSSKRGQLQISTTAAQIAKVMEEVSAIHTS	900
DB	841	LDSSRSSEKDRSLERERIGLGNYPHATENPCTSSKRGQLQISTTAAQIAKVMEEVSAIHTS	900
QY	901	QEDRSSGSGTTELHCVTDERNALRRSSAAHTSNFTKSENSNRTCSMPYAKLEYKRSS	960
DB	901	QEDRSSGSGTTELHCVTDERNALRRSSAAHTSNFTKSENSNRTCSMPYAKLEYKRSS	960
QY	961	NDLSNSVSNDCYKRGQMKPSIESYSEDDSKFCYQYQPADLAHKIHSANHMDNDGE	1020
DB	961	NDLSNSVSNDCYKRGQMKPSIESYSEDDSKFCYQYQPADLAHKIHSANHMDNDGE	1020
QY	1021	LDTPINYSKYSDEQLNSGRSPSONERWARPKHIIIEDEIKQSQORQSRNOSTTYPVYTE	1080
DB	1021	LDTPINYSKYSDEQLNSGRSPSONERWARPKHIIIEDEIKQSQORQSRNOSTTYPVYTE	1080
QY	1081	STDDXHLKXFPHQEQECVSPYRSGANGSETNVSNGHGINQVNSQSLCEDDYEDDKP	1140
DB	1081	STDDXHLKXFPHQEQECVSPYRSGANGSETNVSNGHGINQVNSQSLCEDDYEDDKP	1140
QY	1141	TNYSRYSEEEQHEEERPTNYSIKYNEEKRVDPQIDYSLKYATDIPSSQKQSFSSKS	1200
DB	1141	TNYSRYSEEEQHEEERPTNYSIKYNEEKRVDPQIDYSLKYATDIPSSQKQSFSSKS	1200
QY	1201	SSGQSKTEHSSSSSENTSTPSSNAKQNLHPSASQSGQPOKAATCKVSSINQETIQ	1260
DB	1201	SSGQSKTEHSSSSSENTSTPSSNAKQNLHPSASQSGQPOKAATCKVSSINQETIQ	1260
QY	1261	TYCVEDTPICFSRCSLSSLSAEDEICNQTQEADSANTLQIAEIKGIGTSAEDPV	1320
DB	1261	TYCVEDTPICFSRCSLSSLSAEDEICNQTQEADSANTLQIAEIKGIGTSAEDPV	1320
QY	1321	SEVPASVQHPRTKSRLOGSSLSSESARHKAVERPSSGAKSKSGCACTPKSPHYVQET	1380
DB	1321	SEVPASVQHPRTKSRLOGSSLSSESARHKAVERPSSGAKSKSGCACTPKSPHYVQET	1380
QY	1381	PLMFRCSTSVSLDSFESRSIASSVQSEPCSGMVGIIISPSDLDPSCQOTMPPSRSTPP	1440
DB	1381	PLMFRCSTSVSLDSFESRSIASSVQSEPCSGMVGIIISPSDLDPSCQOTMPPSRSTPP	1440
QY	1441	PPQTAQTKREVPKNAKPTASKRESGPKQAANVAQVQLPADTLLHFATSTPDGF	1500
DB	1441	PPQTAQTKREVPKNAKPTASKRESGPKQAANVAQVQLPADTLLHFATSTPDGF	1500
QY	1501	SCSSLSLSALSDDEPIQKDELIRIMPPVQENDNGNETESQPKESNENQEKAEKTIIDSE	1560
DB	1501	SCSSLSLSALSDDEPIQKDELIRIMPPVQENDNGNETESQPKESNENQEKAEKTIIDSE	1560
QY	1561	KDLDSDSDDDIIELEECIIISAMPTKSRKQKPAQATASKLPPPVARKPSOLPYKLLPS	1620
DB	1561	KDLDSDSDDDIIELEECIIISAMPTKSRKQKPAQATASKLPPPVARKPSOLPYKLLPS	1620
QY	1621	QNRLOPQKHVSFTPGDDMPRVYCVGEGTPIFSTATSLDITIESPPNELAAGEVGRGAQ	1680
DB	1621	QNRLOPQKHVSFTPGDDMPRVYCVGEGTPIFSTATSLDITIESPPNELAAGEVGRGAQ	1680



QY 1681 SCEPEKRDITPEGRSTDEAOGKTSVYTIPELDNKAEBGDILAEACINSAMPKCKSHK 1740  
Db 1681 SCEPEKRDITPEGRSTDEAOGKTSVYTIPELDNKAEBGDILAEACINSAMPKCKSHK 1740  
QY 1741 FRVKIMDQVQASASSAPNKNQLDGKKKPTSPVKPIPQNTYRTRVRKNADSKNIN 1800  
Db 1741 FRVKIMDQVQASASSAPNKNQLDGKKKPTSPVKPIPQNTYRTRVRKNADSKNIN 1800  
QY 1801 AERFSDNKKQKQKNNKNDPNDKLNEDRVRGSAFDSPPHYTIEGTPTCFPSND 1860  
Db 1801 AERFSDNKKQKQKNNKNDPNDKLNEDRVRGSAFDSPPHYTIEGTPTCFPSND 1860  
QY 1861 SLSSLDFFDDDDVLDLREKAEKAKENKESAKVTSHLTNSQOSANKTOAIKQPINR 1920  
Db 1861 SLSSLDFFDDDDVLDLREKAEKAKENKESAKVTSHLTNSQOSANKTOAIKQPINR 1920  
QY 1921 GQPKILOKQSTFPOSSKDIIPRGAADEKONFAIENTPVCFSNSSLSDIDQNN 1980  
Db 1921 GQPKILOKQSTFPOSSKDIIPRGAADEKONFAIENTPVCFSNSSLSDIDQNN 1980  
QY 1981 NKNEPIKETEPDPSQGPSPKQASGYAPKSPHVEDTVCFSRNSLSLSIDSEDDLLQ 2040  
Db 1981 NKNEPIKETEPDPSQGPSPKQASGYAPKSPHVEDTVCFSRNSLSLSIDSEDDLLQ 2040  
QY 2041 ECISAMPKKKPSRLKGDNEKXSPRNMGGILGEDTLTKDQIQRPDSEHGLSPDSENF 2100  
Db 2041 ECISAMPKKKPSRLKGDNEKXSPRNMGGILGEDTLTKDQIQRPDSEHGLSPDSENF 2100  
QY 2101 WKAIQEGANSIVSSLHQAACLRQASSDSILSLKSGISLSPHLPDQEEKPT 2160  
Db 2101 WKAIQEGANSIVSSLHQAACLRQASSDSILSLKSGISLSPHLPDQEEKPT 2160  
QY 2161 SNKGRILKPKGKSTLETKIESEKGIKGGKVKYKSLITGVRNSNLSISGQMQLQAN 2220  
Db 2161 SNKGRILKPKGKSTLETKIESEKGIKGGKVKYKSLITGVRNSNLSISGQMQLQAN 2220  
QY 2221 MPSISGRMTIHPGVRNSSSSTSPVSKGPPKTPASKSPSEGGTATTSRGAKPVS 2280  
Db 2221 MPSISGRMTIHPGVRNSSSSTSPVSKGPPKTPASKSPSEGGTATTSRGAKPVS 2280  
QY 2281 ELSPVAROTSLQGGSKAPSRSGSDSTPSRPAQOPLSRPIQSPGRNISIPGRNISPPN 2340  
Db 2281 ELSPVAROTSLQGGSKAPSRSGSDSTPSRPAQOPLSRPIQSPGRNISIPGRNISPPN 2340  
QY 2341 KLSQLPRTSSPTATKSGSGKMYTSPGRQMSQONLTQTLGSKNASSIPRESASKG 2400  
Db 2341 KLSQLPRTSSPTATKSGSGKMYTSPGRQMSQONLTQTLGSKNASSIPRESASKG 2400  
QY 2401 LNMWNGCANKKVELSRMSSSTKSGSDRSERPVLRQSTFIKEAPSPILRRKLESA 2460  
Db 2401 LNMWNGCANKKVELSRMSSSTKSGSDRSERPVLRQSTFIKEAPSPILRRKLESA 2460  
QY 2461 SPESLSPSSRPASPTRSQAQTPVLSPLPDMSLTHSSVQAGGWRKLPNLSPTIEYNDG 2520  
Db 2461 SPESLSPSSRPASPTRSQAQTPVLSPLPDMSLTHSSVQAGGWRKLPNLSPTIEYNDG 2520  
QY 2521 RPAKHDIARSSESPRLPINRGTWKREHSHSSLPVSTWRTGSSSSILSASSES 2580  
Db 2521 RPAKHDIARSSESPRLPINRGTWKREHSHSSLPVSTWRTGSSSSILSASSES 2580  
QY 2581 SEKAKSEDEKHVNSISGTQKSENOVSAGTWRKIKENEFPSTNSTQTSVSGATNGAES 2640  
Db 2581 SEKAKSEDEKHVNSISGTQKSENOVSAGTWRKIKENEFPSTNSTQTSVSGATNGAES 2640  
QY 2641 KTLIYQMAFPAVKTEDEVVWRIEDCPINNPRSGRSTGNTPPVIDSVSEKAPNPKDSKN 2700  
Db 2641 KTLIYQMAFPAVKTEDEVVWRIEDCPINNPRSGRSTGNTPPVIDSVSEKAPNPKDSKN 2700  
QY 2701 QAKQNVGSGVPMRTVGLNRLTSLFIQVADPQKTEIKPGQNNPVPVSETNESPIVERT 2760  
Db 2701 QAKQNVGSGVPMRTVGLNRLTSLFIQVADPQKTEIKPGQNNPVPVSETNESPIVERT 2760

QY 2761 PFSSSSSKHSSPSGTVAARVTPFNPNPSPRKSSAUSTSARPSQIPTPVNNNTKKRDSKT 2820  
Db 2761 PFSSSSSKHSSPSGTVAARVTPFNPNPSPRKSSAUSTSARPSQIPTPVNNNTKKRDSKT 2820  
QY 2821 DSTESSGTQSPKHSGLSYLTVS 2843  
Db 2821 DSTESSGTQSPKHSGLSYLTVS 2843

RESULT 8  
US-08-450-582-2  
; Sequence 2, Application US/08450582  
; Patent No. 6114124  
; GENERAL INFORMATION:  
; APPLICANT: ALBERTSEN, HANS  
; APPLICANT: ANAND, RAKESH  
; APPLICANT: CARLSON, MARY  
; APPLICANT: GRODEN, JOANNA  
; APPLICANT: HEDGE, PHILIP J.  
; APPLICANT: JOSLYN, GEOFF  
; APPLICANT: KINZLER, KENNETH  
; APPLICANT: MARKHAM, ALEXANDER F.  
; APPLICANT: NAKAMURA, YUSUKE  
; APPLICANT: THLIVERIS, ANDREW  
; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC  
; TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS  
; NUMBER OF SEQUENCES: 102  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Banner & Witcoff, Ltd.  
; STREET: 1001 G Street, NW  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20001-4598  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/450,582  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/452,655  
; FILING DATE: 25-MAY-1995  
; APPLICATION NUMBER: US 08/289,548  
; FILING DATE: 12-AUG-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/741,940  
; FILING DATE: 08-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kagan, Sarah A.  
; REGISTRATION NUMBER: 32,141  
; REFERENCE/DOCKET NUMBER: 1107.49964  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-508-9100  
; TELEFAX: 202-508-9239  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2843 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-450-582-2

Query Match 99.7%; Score 14533; DB 3; Length 2843;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2836; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 MAAASYDOLLKQVLEALKNWNSLRQELDNSNHLTKLETEASNMKEVLKOLQGSIEDAM 60  
Db 1 MAAASYDOLLKQVLEALKNWNSLRQELDNSNHLTKLETEASNMKEVLKOLQGSIEDAM 60

QY 61 ASSQIDLLERLKEINLNDNFPGVKLRKMSLRYSRGSGSVSSRGSGCSPVPMGSPFR 120  
 DB 61 ASSQIDLLERLKEINLNDNFPGVKLRKMSLRYSRGSGSVSSRGSGCSPVPMGSPFR 120  
 QY 121 RGFGNGRESTGYLEBELEKRSLLADLDEEKEKOWYQAOLQNLKRGIDSLPLTENFSL 180  
 DB 121 RGFGNGRESTGYLEBELEKRSLLADLDEEKEKOWYQAOLQNLKRGIDSLPLTENFSL 180  
 QY 181 QTLTRQLEYEARQIRVAMEEOLGTQDMEXKAQRRIARIQOIEKIDILRIROLLSQAT 240  
 DB 181 QTMTRQLEYEARQIRVAMEEOLGTQDMEXKAQRRIARIQOIEKIDILRIROLLSQAT 240  
 QY 241 EABRSSQKHETGSHDAERNEGOGVGEINMATSGNGQGGSTTRMDHETASVLSSTHSA 300  
 DB 241 EABRSSQKHETGSHDAERNEGOGVGEINMATSGNGQGGSTTRMDHETASVLSSTHSA 300  
 QY 301 PRRLTSHLGTKEVNYSLLSMLGTHDXDMSRTLILAMSSQDSCISMRQSGCLPLLIQLL 360  
 DB 301 PRRLTSHLGTKEVNYSLLSMLGTHDXDMSRTLILAMSSQDSCISMRQSGCLPLLIQLL 360  
 QY 361 HGNDKDSVLLGNSRGSKEARASAAALHNIHSGOPDDKRGRRBIRVLHLEQIRAYCETC 420  
 DB 361 HGNDKDSVLLGNSRGSKEARASAAALHNIHSGOPDDKRGRRBIRVLHLEQIRAYCETC 420  
 QY 421 WEQEAHEPGMDOKNMPAPVHOICPAVCVLMLKLSFDEEHRHAMELGGLOAIAELLO 480  
 DB 421 WEQEAHEPGMDOKNMPAPVHOICPAVCVLMLKLSFDEEHRHAMELGGLOAIAELLO 480  
 QY 481 VDCMYGLTNDHYGITTIRRYAGMALTNLTFGDVANKATLCMSKGMCMALVAQLKSEEDL 540  
 DB 481 VDCMYGLTNDHYGITTIRRYAGMALTNLTFGDVANKATLCMSKGMCMALVAQLKSEEDL 540  
 QY 541 QOVTASVLRLNLSWADVNSKTLREVGSVKALMECALEVKESTLKSVALNLSAHC 600  
 DB 541 QOVTASVLRLNLSWADVNSKTLREVGSVKALMECALEVKESTLKSVALNLSAHC 600  
 QY 601 ENKADI CAVDGALAFVGLTYRSQTNLTALIESGGILRNVSLLATNEDHRIQLRENN 660  
 DB 601 ENKADI CAVDGALAFVGLTYRSQTNLTALIESGGILRNVSLLATNEDHRIQLRENN 660  
 QY 661 CLQTLLOHLKSHSLTIVSNACGTLNLSARNPKDQEAALMDMGAVSMLKNIHSHKHMIAM 720  
 DB 661 CLQTLLOHLKSHSLTIVSNACGTLNLSARNPKDQEAALMDMGAVSMLKNIHSHKHMIAM 720  
 QY 721 GSAAALRLNLRNPAKYKDANIMSPGSLPSLHVKKOKALEAELDAHLSETFDNLDLS 780  
 DB 721 GSAAALRLNLRNPAKYKDANIMSPGSLPSLHVKKOKALEAELDAHLSETFDNLDLS 780  
 QY 781 PKASHRSKORHKQSLGYDVFDTNRHDDNRSDNFNTGNMTVLPYLNTTVLPSSSSRGS 840  
 DB 781 PKASHRSKORHKQSLGYDVFDTNRHDDNRSDNFNTGNMTVLPYLNTTVLPSSSSRGS 840  
 QY 841 LDSRSSEKORSLERGIGICNVHPATENPCTSKRGLOISTTAAQIAKVMESVAIHTS 900  
 DB 841 LDSRSSEKORSLERGIGICNVHPATENPCTSKRGLOISTTAAQIAKVMESVAIHTS 900  
 QY 901 QEDRSSGTTTELHCVTDERNALRSSAAHSTNFTNFTKSENSNRCTCSMPYAKLEYKRS 960  
 DB 901 QEDRSSGTTTELHCVTDERNALRSSAAHSTNFTNFTKSENSNRCTCSMPYAKLEYKRS 960  
 QY 961 NDSLNSVSSNDGYKRGOMKPSIESYEDDESKEFCYQYPADIAHKIHSANHMDNDGE 1020  
 DB 961 NDSLNSVSSNDGYKRGOMKPSIESYEDDESKEFCYQYPADIAHKIHSANHMDNDGE 1020  
 QY 1021 LDTPINTSLKYSDQLNSGRQSPQNERWARPKHIIEDKIQEQRQSRNQSTTYPVYTE 1080  
 DB 1021 LDTPINTSLKYSDQLNSGRQSPQNERWARPKHIIEDKIQEQRQSRNQSTTYPVYTE 1080  
 QY 1081 STDDKHLKFPQHFQOQCVSPYRSRGANGSETNRVGSNHGINQVNSQSLCQEDDYEDDXP 1140  
 DB 1081 STDDKHLKFPQHFQOQCVSPYRSRGANGSETNRVGSNHGINQVNSQSLCQEDDYEDDXP 1140

QY 1141 TNSYERYEEHQHEEERPTNYSIKYNEKHHVQOPIDYSLKYATDIPSSQKQSFSSKS 1200  
 DB 1141 TNSYERYEEHQHEEERPTNYSIKYNEKHHVQOPIDYSLKYATDIPSSQKQSFSSKS 1200  
 QY 1201 SSGQSSKTEHWSSESSENTSTPSSNAKQNLHPSSAQRSGPOKAAATCKVSSINQTIQ 1260  
 DB 1201 SSGQSSKTEHWSSESSENTSTPSSNAKQNLHPSSAQRSGPOKAAATCKVSSINQTIQ 1260  
 QY 1261 TYCVEDTPICFRSRCSLSLSAEDICNCTTQFADSANTLQIAEIKGKTGTRSAEDPV 1320  
 DB 1261 TYCVEDTPICFRSRCSLSLSAEDICNCTTQFADSANTLQIAEIKGKTGTRSAEDPV 1320  
 QY 1321 SEVPASVQHPRTKSRLOGSSLSSESARHKAVERPSGAKSPSKSGAOTPKSPHYVQET 1380  
 DB 1321 SEVPASVQHPRTKSRLOGSSLSSESARHKAVERPSGAKSPSKSGAOTPKSPHYVQET 1380  
 QY 1381 PLMFRCCTSVSLSLDSFESRSIASSVQSEPCSGMVSGIISPSDLPSDQTMPPSRKSTPP 1440  
 DB 1381 PLMFRCCTSVSLSLDSFESRSIASSVQSEPCSGMVSGIISPSDLPSDQTMPPSRKSTPP 1440  
 QY 1441 PPQTAQTKREVPPKNKAPTABKRESGPKQAANAARVQVLPDADTLILHFAESTPDGF 1500  
 DB 1441 PPQTAQTKREVPPKNKAPTABKRESGPKQAANAARVQVLPDADTLILHFAESTPDGF 1500  
 QY 1501 SCSSLSLSLSDDEPFIQKDVLRIMPVQENDNGNETESEOPKESNENQEKAEKTIIDSE 1560  
 DB 1501 SCSSLSLSLSDDEPFIQKDVLRIMPVQENDNGNETESEOPKESNENQEKAEKTIIDSE 1560  
 QY 1561 KOLLDDSDDDDIIELEECIISAMPTKSRKOKPAQATASKLPPPVARKPSOLPVYKLLPS 1620  
 DB 1561 KOLLDDSDDDDIIELEECIISAMPTKSRKOKPAQATASKLPPPVARKPSOLPVYKLLPS 1620  
 QY 1621 QNRLOPQKHVSFTGDDMPRVYCVETGPINFSTATSLSDLTIESPPNELAAGEVGRGAQ 1680  
 DB 1621 QNRLOPQKHVSFTGDDMPRVYCVETGPINFSTATSLSDLTIESPPNELAAGEVGRGAQ 1680  
 QY 1681 SGEPEKRTDITTEGRSTDEAOGGKTSVTIPELDDNKAEEGDILAEICINSAMPKGSHPK 1740  
 DB 1681 SGEPEKRTDITTEGRSTDEAOGGKTSVTIPELDDNKAEEGDILAEICINSAMPKGSHPK 1740  
 QY 1741 FRVKIMDOVQOASASSAPNKNQIDGKKKXFTSPVXPIQONTYRTRVRKADSKNIN 1800  
 DB 1741 FRVKIMDOVQOASASSAPNKNQIDGKKKXFTSPVXPIQONTYRTRVRKADSKNIN 1800  
 QY 1801 AERFSDNKNDSKKQNLKNNKDFNDKLPNNEDRVGSGFAFDSPHHYTPIEGTPYCFSRND 1860  
 DB 1801 AERFSDNKNDSKKQNLKNNKDFNDKLPNNEDRVGSGFAFDSPHHYTPIEGTPYCFSRND 1860  
 QY 1861 SLSSLDPDDDVDVLSREKAEILRKAKENKESPAKTSHTELTSNOQSANKTOAIKQPINR 1920  
 DB 1861 SLSSLDPDDDVDVLSREKAEILRKAKENKESPAKTSHTELTSNOQSANKTOAIKQPINR 1920  
 QY 1921 GQPKFLOKQSTFPQSSKDI PDRGAATDEKLQNFIENTPVCFSHNSLSLSLSDIDQENN 1980  
 DB 1921 GQPKFLOKQSTFPQSSKDI PDRGAATDEKLQNFIENTPVCFSHNSLSLSLSDIDQENN 1980  
 QY 1981 NKNEPIKETEPPOSQSGEPSPKQASGVAPKSFHVEDTVPFCSRNSLSLSLSDIDEDLLQ 2040  
 DB 1981 NKNEPIKETEPPOSQSGEPSPKQASGVAPKSFHVEDTVPFCSRNSLSLSLSDIDEDLLQ 2040  
 QY 2041 ECISAMPKKKKSRLLKGDNEKHSFRNMGGLGDLTLDLKDIOKPDSEHGLSPDSNFD 2100  
 DB 2041 ECISAMPKKKKSRLLKGDNEKHSFRNMGGLGDLTLDLKDIOKPDSEHGLSPDSNFD 2100  
 QY 2101 WKAIQEGANSIVSSLHQAAAAACLSRQASSDSILSKSGISLGSFPFLTTPDQEEKPFT 2160  
 DB 2101 WKAIQEGANSIVSSLHQAAAAACLSRQASSDSILSKSGISLGSFPFLTTPDQEEKPFT 2160  
 QY 2161 SNKGRIILKPEKSTLETETKTESKGTGGKVKYKSLITGKVSNSSEISQOMKQIPLOAN 2220  
 DB 2161 SNKGRIILKPEKSTLETETKTESKGTGGKVKYKSLITGKVSNSSEISQOMKQIPLOAN 2220  
 QY 2221 MFSIRSGRTMIHIFGVRNRSSTSVSVKKGKPLKTPASKSPSEGTATTSPRGAKPSVKS 2280

Db 2221 MFESISGRMTIHIPGVRNSSSTSPVSKGPKLPKTPASKSPSEGTATTSPRGAKPSVKS 2280  
Qy 2281 ELSPVARTSQIGSGSKAPSRGSRDSTPSRPAQOPLSRPIQSPGRNIGSPGRNIGSPFN 2340  
Db 2281 ELSPVARTSQIGSGSKAPSRGSRDSTPSRPAQOPLSRPIQSPGRNIGSPGRNIGSPFN 2340  
Qy 2341 KLSQLPRTSSPSTASTKSSGSKMYSTSPGRMSQOQLTKQTLGSKNASSIPRSESASKG 2400  
Db 2341 KLSQLPRTSSPSTASTKSSGSKMYSTSPGRMSQOQLTKQTLGSKNASSIPRSESASKG 2400  
Qy 2401 LNMQNGGANKKVELSRMSSSTKSSGSDRSERPVLRQSTFIKEAPSPILRRKLEESA 2460  
Db 2401 LNMQNGGANKKVELSRMSSSTKSSGSDRSERPVLRQSTFIKEAPSPILRRKLEESA 2460  
Qy 2461 SFESLSPSRPASPTRSQAOTPVLSPLDMSLSTHSSVQAGGWRKLPNLSPTIYNDG 2520  
Db 2461 SFESLSPSRPASPTRSQAOTPVLSPLDMSLSTHSSVQAGGWRKLPNLSPTIYNDG 2520  
Qy 2521 RPAKHDIARSHSESPRLPNRSGTWKREHSHKSSSLPRVSTWRTGSSSILSASSES 2580  
Db 2521 RPAKHDIARSHSESPRLPNRSGTWKREHSHKSSSLPRVSTWRTGSSSILSASSES 2580  
Qy 2581 SEKAKSEDEKHVNSISGTRKQSKENQVSAKGTWRKIKENEFSPNTSQTSSGATNGAES 2640  
Db 2581 SEKAKSEDEKHVNSISGTRKQSKENQVSAKGTWRKIKENEFSPNTSQTSSGATNGAES 2640  
Qy 2641 KTLIYQMAPAVSKTEDEVWRIEDCPINNPRSGRSTGNTPPVIDSVSEKAPNKKSDN 2700  
Db 2641 KTLIYQMAPAVSKTEDEVWRIEDCPINNPRSGRSTGNTPPVIDSVSEKAPNKKSDN 2700  
Qy 2701 QAKQNVGSGVPMRTVGLNRLTSTIOVDAPQKTEIKPGQNNPVPVSETNESPIVBT 2760  
Db 2701 QAKQNVGSGVPMRTVGLNRLTSTIOVDAPQKTEIKPGQNNPVPVSETNESPIVBT 2760  
Qy 2761 PFSSSSSSKHSPSGTVAARVTPFNYPNPSRKSSADSTSARPSQIPTPVNNTKKRDSKT 2820  
Db 2761 PFSSSSSSKHSPSGTVAARVTPFNYPNPSRKSSADSTSARPSQIPTPVNNTKKRDSKT 2820  
Qy 2821 DSTESSGTSQPKRHSGSYLVTSV 2843  
Db 2821 DSTESSGTSQPKRHSGSYLVTSV 2843

RESULT 9

US-08-450-582-7  
; Sequence 7, Application US/08450582.  
; Patent No. 6114124  
; GENERAL INFORMATION:  
; APPLICANT: ALBERTSEN, HANS  
; APPLICANT: ANAND, RAKESH  
; APPLICANT: CARLSON, MARY  
; APPLICANT: GRODEN, JOANNA  
; APPLICANT: HEDGE, PHILIP J.  
; APPLICANT: JOSLYN, GEOFF  
; APPLICANT: KINZLER, KENNETH  
; APPLICANT: MARKHAM, ALEXANDER F.  
; APPLICANT: NAKAMURA, YUSUKE  
; APPLICANT: THLIVERIS, ANDREW  
; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC  
; TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS  
; NUMBER OF SEQUENCES: 102  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Banner & Witcoff, Ltd.  
; STREET: 1001 G Street, NW  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20001-4598  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/450,582  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/452,655  
FILING DATE: 25-MAY-1995  
APPLICATION NUMBER: US 08/289,548  
FILING DATE: 12-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/741,940  
FILING DATE: 08-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: kagan, Sarah A.  
REGISTRATION NUMBER: 32,141  
REFERENCE/DOCKET NUMBER: 1107.49964  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-508-9100  
TELEFAX: 202-508-9299  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2843 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHEICAL: YES  
ANTI-SENSE: NO  
US-08-450-582-7

Query Match 99.7%; Score 14533; DB 3; Length 2843;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2836; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
Qy 1 MAAASYDQLLKQVEALKWENSINLROLEDNSNHLTKLETEASNMKEVLKOLQSIIDEAM 60  
Db 1 MAAASYDQLLKQVEALKWENSINLROLEDNSNHLTKLETEASNMKEVLKOLQSIIDEAM 60  
Qy 61 ASSGOIDLLRLKELNLDSSNFPQVKLRSMKLSRSGSREGSVSSSGESCPVPMGSFPR 120  
Db 61 ASSGOIDLLRLKELNLDSSNFPQVKLRSMKLSRSGSREGSVSSSGESCPVPMGSFPR 120  
Qy 121 RGFVNGSRESTGYLELEKERSILLADLDKEEKEKWYQAOLQNLTKRIDLPLTENFSL 180  
Db 121 RGFVNGSRESTGYLELEKERSILLADLDKEEKEKWYQAOLQNLTKRIDLPLTENFSL 180  
Qy 181 QTDLTREOLEYEAQIRVAVEEQLGTCODMEKRAQRIARIQIIEKDIRIOLLOSCAT 240  
Db 181 QTDLTREOLEYEAQIRVAVEEQLGTCODMEKRAQRIARIQIIEKDIRIOLLOSCAT 240  
Qy 241 EAERSQNKHETGSHDAERQNEGGVGEINMATSGNGQGGSTTRMDHETASVLSSTHSA 300  
Db 241 EAERSQNKHETGSHDAERQNEGGVGEINMATSGNGQGGSTTRMDHETASVLSSTHSA 300  
Qy 301 PRLTSHLGTKEVWYVSLLSMLGTHDKDDMSRILLAMSSQSDSCINRQSGCPLLIQLL 360  
Db 301 PRLTSHLGTKEVWYVSLLSMLGTHDKDDMSRILLAMSSQSDSCINRQSGCPLLIQLL 360  
Qy 361 HGNDKSVLLGNSRGSKEARASAAALHNIHSQPDQKRRRIRVILHLEQIRAYCETC 420  
Db 361 HGNDKSVLLGNSRGSKEARASAAALHNIHSQPDQKRRRIRVILHLEQIRAYCETC 420  
Qy 421 WEMQEAHEPGMDQKMPAPVSHQICPAVCVLMKLSFDEEHRHAMNELGGLQAIALLQ 480  
Db 421 WEMQEAHEPGMDQKMPAPVSHQICPAVCVLMKLSFDEEHRHAMNELGGLQAIALLQ 480  
Qy 481 VDCMYGLTNDHYSITLRRYAGMALTNLTFGDVANKATLCSMKGCRALVAQKSEEDL 540  
Db 481 VDCMYGLTNDHYSITLRRYAGMALTNLTFGDVANKATLCSMKGCRALVAQKSEEDL 540  
Qy 541 QQVIASVLRNLSWRADYNSKTLREVSGSVKALMECALEVKESTLKSIVLSALWNLSAHT 600  
Db 541 QQVIASVLRNLSWRADYNSKTLREVSGSVKALMECALEVKESTLKSIVLSALWNLSAHT 600

Db 541 QOVIASVLRNLSMRADYNKKTREVGSVKALMECALEVKESTLKSIVLSALWNLSAHCT 600  
Qy 601 ENKADICAVDICALAFVGLTYRSQTNLTALIESGGGILRNVSLLIATNEDHRQILRENN 660  
Db 601 ENKADICAVDICALAFVGLTYRSQTNLTALIESGGGILRNVSLLIATNEDHRQILRENN 660  
Qy 661 CLOTLLQHLKSHSLTIVSNACGLTWNLSARNPKDQOEALWDMGAVSMKNIHISKHOMIAM 720  
Db 661 CLOTLLQHLKSHSLTIVSNACGLTWNLSARNPKDQOEALWDMGAVSMKNIHISKHOMIAM 720  
Qy 721 GSAALRNLMANRPAYKNDANIMSPGSSLPKLVKOKALEAELDAQHSETFDNIDNLS 780  
Db 721 GSAALRNLMANRPAYKNDANIMSPGSSLPKLVKOKALEAELDAQHSETFDNIDNLS 780  
Qy 781 PKASHRSKQRHKQSLYGDYVFDNRHDDNRSDFNFTGNMTVLPSPYLNNTVLPSSSSSRGS 840  
Db 781 PKASHRSKQRHKQSLYGDYVFDNRHDDNRSDFNFTGNMTVLPSPYLNNTVLPSSSSSRGS 840  
Qy 841 LDSSRSKORSLRERGIGLGNHPATENPCTGSKGLQISTTAAQIAKYMEEVSAIHTS 900  
Db 841 LDSSRSKORSLRERGIGLGNHPATENPCTGSKGLQISTTAAQIAKYMEEVSAIHTS 900  
Qy 901 QEDRSSGTTTELHCVTDERNALRSSAAHSTHNTYNTFKSENRTCSMPYAKLEYKRSS 960  
Db 901 QEDRSSGTTTELHCVTDERNALRSSAAHSTHNTYNTFKSENRTCSMPYAKLEYKRSS 960  
Qy 961 NDSLNSVSSNDGCKRQMKPESIBSYEDDESKEFCSCYQYPADLAHKSANHMDNDGE 1020  
Db 961 NDSLNSVSSNDGCKRQMKPESIBSYEDDESKEFCSCYQYPADLAHKSANHMDNDGE 1020  
Qy 1021 LDTPINTSLKYSDQLNSGRQSPONERWARPKHIIIEDEIKQSEQRQSRNQSTTYPVYTE 1080  
Db 1021 LDTPINTSLKYSDQLNSGRQSPONERWARPKHIIIEDEIKQSEQRQSRNQSTTYPVYTE 1080  
Qy 1081 STDDKHLKFPQHFQOQCVSPYRSRGANGSETNRVGSNHGINQVSGSLCOEDDYEDDKP 1140  
Db 1081 STDDKHLKFPQHFQOQCVSPYRSRGANGSETNRVGSNHGINQVSGSLCOEDDYEDDKP 1140  
Qy 1141 TNYSERVSEBQHEERPTNYSIKYNEBKHVDPIDYSLKYATDIPSSOKQSFPSKS 1200  
Db 1141 TNYSERVSEBQHEERPTNYSIKYNEBKHVDPIDYSLKYATDIPSSOKQSFPSKS 1200  
Qy 1201 SSGQSKTEHWSSESSTSTPSSNAKQNLHPSSAQSRGQPKAATCKVSSINQETIQ 1260  
Db 1201 SSGQSKTEHWSSESSTSTPSSNAKQNLHPSSAQSRGQPKAATCKVSSINQETIQ 1260  
Qy 1261 TYCVEDTPICFSRCSLSLSAABDEIGCQNTQOEADSAANTLOIAETIKGIGTRSAEDPV 1320  
Db 1261 TYCVEDTPICFSRCSLSLSAABDEIGCQNTQOEADSAANTLOIAETIKGIGTRSAEDPV 1320  
Qy 1321 SEVPAVQHPRTKSRRLQSSLSSESARHKAVERFPSSGAKSFSGAQTPKSPPEHYVQET 1380  
Db 1321 SEVPAVQHPRTKSRRLQSSLSSESARHKAVERFPSSGAKSFSGAQTPKSPPEHYVQET 1380  
Qy 1381 PLMFSRCTSVSLDSFESRSIASSVQSEPCSGMWGIIISPDLPDSGQTPMPSRSKTPP 1440  
Db 1381 PLMFSRCTSVSLDSFESRSIASSVQSEPCSGMWGIIISPDLPDSGQTPMPSRSKTPP 1440  
Qy 1441 PPPQTAQTKREVKNKAPTAEKRESGPKQAANAAVQVQVLPDADTLLHFAFATESPDGF 1500  
Db 1441 PPPQTAQTKREVKNKAPTAEKRESGPKQAANAAVQVQVLPDADTLLHFAFATESPDGF 1500  
Qy 1501 SCSSLSALSLEDBPFIQKVDELRLMPVQVNDNGNETESOPKESNENQKEAEKTI DSE 1560  
Db 1501 SCSSLSALSLEDBPFIQKVDELRLMPVQVNDNGNETESOPKESNENQKEAEKTI DSE 1560  
Qy 1561 KDLDDSDDDDDIEILEBCCIISAMPTKSRKGPQAOTASKLPPPVARKPSQLPVYKLLPS 1620  
Db 1561 KDLDDSDDDDDIEILEBCCIISAMPTKSRKGPQAOTASKLPPPVARKPSQLPVYKLLPS 1620  
Qy 1621 QNRLOPQKHVSFTPGDDMPRVYCVGEGTPIINFSTATSLSDLTIESPPNELAAGSVRGGAQ 1680  
Db 1621 QNRLOPQKHVSFTPGDDMPRVYCVGEGTPIINFSTATSLSDLTIESPPNELAAGSVRGGAQ 1680

Qy 1681 SGEFEKRDITPTEGRSTDEAOGGKTSSVTIPELDDNKAEBGDILAEICINSAMPKGSHPK 1740  
Db 1681 SGEFEKRDITPTEGRSTDEAOGGKTSSVTIPELDDNKAEBGDILAEICINSAMPKGSHPK 1740  
Qy 1741 FRVKIMDOVOOASASSAPNKNOLDGKKKPTSPVKPIPONTREYRTRVRKNADSKNLN 1800  
Db 1741 FRVKIMDOVOOASASSAPNKNOLDGKKKPTSPVKPIPONTREYRTRVRKNADSKNLN 1800  
Qy 1801 AERVFSDNKDSKKQNLKNKNSKDFNDKLPNNEDRVGSAFDFSPHHYTPIEGTYPYCFSRND 1860  
Db 1801 AERVFSDNKDSKKQNLKNKNSKDFNDKLPNNEDRVGSAFDFSPHHYTPIEGTYPYCFSRND 1860  
Qy 1861 SLSLDFDDDDVDLSREKAEIRKAKENKESFAKYTSHTELTSNQQSANKTQAIKQPINR 1920  
Db 1861 SLSLDFDDDDVDLSREKAEIRKAKENKESFAKYTSHTELTSNQQSANKTQAIKQPINR 1920  
Qy 1921 GQPKPILQKQSTFPQSSKDIIPDRGAATDEKLQNFPAIENTPVCFSHNSLSLSLSDIDQENN 1980  
Db 1921 GQPKPILQKQSTFPQSSKDIIPDRGAATDEKLQNFPAIENTPVCFSHNSLSLSLSDIDQENN 1980  
Qy 1981 NKEBPIKETPPOSGEPSPKQASGYAPKSFHYEDTVPFCSRNSLSLSLSDIDQENN 2040  
Db 1981 NKEBPIKETPPOSGEPSPKQASGYAPKSFHYEDTVPFCSRNSLSLSLSDIDQENN 2040  
Qy 2041 ECISAMPKKKPRLKGDNEKHSFRNNGGILGBDLTLDKDIORPSEHGLSPDSNF 2100  
Db 2041 ECISAMPKKKPRLKGDNEKHSFRNNGGILGBDLTLDKDIORPSEHGLSPDSNF 2100  
Qy 2101 WKAIQEGANSIVSLHQAAAAACLSRQASDSDSILSKSGISILGSPFHLTPDOEKPF 2160  
Db 2101 WKAIQEGANSIVSLHQAAAAACLSRQASDSDSILSKSGISILGSPFHLTPDOEKPF 2160  
Qy 2161 SNKGPRILLKPEKSTLETKTIESKGIKGGKVKYKSLITGKVRNSSEISQMKQPIQAN 2220  
Db 2161 SNKGPRILLKPEKSTLETKTIESKGIKGGKVKYKSLITGKVRNSSEISQMKQPIQAN 2220  
Qy 2221 MFSISRGRTMIHIFGVNRSSSSTSFVSKGPKLPKASPSECOATTTPRGAAPSVKS 2280  
Db 2221 MFSISRGRTMIHIFGVNRSSSSTSFVSKGPKLPKASPSECOATTTPRGAAPSVKS 2280  
Qy 2281 ELSPVARTSOIGSSKAPSRSGRSDTTPSPAQOPLSRPIQSGRNSISPRNGISPPN 2340  
Db 2281 ELSPVARTSOIGSSKAPSRSGRSDTTPSPAQOPLSRPIQSGRNSISPRNGISPPN 2340  
Qy 2341 KLSQLPRTSSPSTASTKSSGSKNSYTSRQMSQOQLTKQTGLSKNASSIPRSESASKG 2400  
Db 2341 KLSQLPRTSSPSTASTKSSGSKNSYTSRQMSQOQLTKQTGLSKNASSIPRSESASKG 2400  
Qy 2401 LNMNNGANKKVELSRMSSTKSSGSDRSEPRVLRQSTFIKEAPSPFLRKLRESA 2460  
Db 2401 LNMNNGANKKVELSRMSSTKSSGSDRSEPRVLRQSTFIKEAPSPFLRKLRESA 2460  
Qy 2461 SPESLSRSPASPTRSQAQTPVLSPSLPDMSLTHSSVQAGGWRKLPNPLSPRIEYNDG 2520  
Db 2461 SPESLSRSPASPTRSQAQTPVLSPSLPDMSLTHSSVQAGGWRKLPNPLSPRIEYNDG 2520  
Qy 2521 RPAKHDIARSHSPSPRLPNNRSGTWKHSKHSLSLPRVSTWRTGSSSSILSASSES 2580  
Db 2521 RPAKHDIARSHSPSPRLPNNRSGTWKHSKHSLSLPRVSTWRTGSSSSILSASSES 2580  
Qy 2581 SEKAKSEDEKVNSTSGTKQSKENQVSAKGTWRKI KENERPSTNSTQTVSSGATNGAES 2640  
Db 2581 SEKAKSEDEKVNSTSGTKQSKENQVSAKGTWRKI KENERPSTNSTQTVSSGATNGAES 2640  
Qy 2641 KTLTYQMAPAVSKTEDVWVRIEDCPINNPRSGRPTGNTPTVIDSVSEKANPNIKDSKDN 2700  
Db 2641 KTLTYQMAPAVSKTEDVWVRIEDCPINNPRSGRPTGNTPTVIDSVSEKANPNIKDSKDN 2700  
Qy 2701 QAKONVGVNGSVPMRTVGLNENLTSFIOVADPDKGTEIKPGQNNPVPVSEINESPIVER 2760  
Db 2701 QAKONVGVNGSVPMRTVGLNENLTSFIOVADPDKGTEIKPGQNNPVPVSEINESPIVER 2760

QY 781 PKASHRSKORHKQSLYGYVDFDTHRHDDNRSDNNTGNTVLSPLYLNTTVLPSSSSSRGS 840  
DB 781 PKASHRSKORHKQSLYGYVDFDTHRHDDNRSDNNTGNTVLSPLYLNTTVLPSSSSSRGS 840  
QY 841 LDSSRSKDRSLERERGLGNYPHATENPGTSSKRGLOISTTAAQIAKWEVSAIHTS 900  
DB 841 LDSSRSKDRSLERERGLGNYPHATENPGTSSKRGLOISTTAAQIAKWEVSAIHTS 900  
QY 901 QEDRSSGTTTLCVTDERNALRSSAAHTSNNTYNTFTKSENSNRTCSMPYAKLEYKRSS 960  
DB 901 QEDRSSGTTTLCVTDERNALRSSAAHTSNNTYNTFTKSENSNRTCSMPYAKLEYKRSS 960  
QY 961 NDSLNSVSSSGYGRGOWKPSIESYSEDDSKSCSYGOYPADIAHKHTSHANHMDNDGE 1020  
DB 961 NDSLNSVSSSGYGRGOWKPSIESYSEDDSKSCSYGOYPADIAHKHTSHANHMDNDGE 1020  
QY 1021 LDTPINYSLKYSDEQLNSGRQSPS QNERWARPKHIIIEIKQSEQRQSRNQSTTYPVYTE 1080  
DB 1021 LDTPINYSLKYSDEQLNSGRQSPS QNERWARPKHIIIEIKQSEQRQSRNQSTTYPVYTE 1080  
QY 1081 STDDKHLKFPQHFQGOECVPSYRSGANGSETNRVGSNHGINONVQSLOCDEDDYEDDKP 1140  
DB 1081 STDDKHLKFPQHFQGOECVPSYRSGANGSETNRVGSNHGINONVQSLOCDEDDYEDDKP 1140  
QY 1141 TNSERYSEEEQHEBEERPTNYSIKYNEEKSHVDOPIDYSLKYATDIPSSQKQSFSEKS 1200  
DB 1141 TNSERYSEEEQHEBEERPTNYSIKYNEEKSHVDOPIDYSLKYATDIPSSQKQSFSEKS 1200  
QY 1201 SSGQSKTEHMSSESENTSTPSNAKRONQLHPSSAQSRSGQPOKAATCKVSSINQETIQ 1260  
DB 1201 SSGQSKTEHMSSESENTSTPSNAKRONQLHPSSAQSRSGQPOKAATCKVSSINQETIQ 1260  
QY 1261 TYCVEDTPICFRCSLSLSAABEIDICNQTQEADSANTLQIAEIKEKIGTSAEDPV 1320  
DB 1261 TYCVEDTPICFRCSLSLSAABEIDICNQTQEADSANTLQIAEIKEKIGTSAEDPV 1320  
QY 1321 SEVPAVSHPRKTSRLQCGSLSSSARHKA VEFSSGAKSPSKGAQTPKPPPEHYVQET 1380  
DB 1321 SEVPAVSHPRKTSRLQCGSLSSSARHKA VEFSSGAKSPSKGAQTPKPPPEHYVQET 1380  
QY 1381 PLMFRCSTSVSLDSFERSSTASSVQSPCGMWVGIIISPDLDPSCQOTMPPRSKTPP 1440  
DB 1381 PLMFRCSTSVSLDSFERSSTASSVQSPCGMWVGIIISPDLDPSCQOTMPPRSKTPP 1440  
QY 1441 PPQTAQTKREVPKNAKPTAEKRESGPKQAAVNAAVQVQVLPDADTLLHFATESTPDGF 1500  
DB 1441 PPQTAQTKREVPKNAKPTAEKRESGPKQAAVNAAVQVQVLPDADTLLHFATESTPDGF 1500  
QY 1501 SCSSLSLSLDEPTIQDVELRIMPVQENDNGNETESBOPKESNENQEKAEKTIIDSE 1560  
DB 1501 SCSSLSLSLDEPTIQDVELRIMPVQENDNGNETESBOPKESNENQEKAEKTIIDSE 1560  
QY 1561 KDLLDSDDDDDTEILBECIIISAMPTKSRKAKKPAQTASKLPPPVARKPSOLPVYKLLPS 1620  
DB 1561 KDLLDSDDDDDTEILBECIIISAMPTKSRKAKKPAQTASKLPPPVARKPSOLPVYKLLPS 1620  
QY 1621 QNRLOPQKHVSFTPGDDMPRVYCVEGTPIINFSTATSLDSTIESPPNELAAGEVGRGAQ 1680  
DB 1621 QNRLOPQKHVSFTPGDDMPRVYCVEGTPIINFSTATSLDSTIESPPNELAAGEVGRGAQ 1680  
QY 1681 SCEPEKRDITPTEGRSTDAQGGKTSVVTIPELDNKAEEGDILAEACINSAMPKGSHPK 1740  
DB 1681 SCEPEKRDITPTEGRSTDAQGGKTSVVTIPELDNKAEEGDILAEACINSAMPKGSHPK 1740  
QY 1741 FRVKKIMDOVQOASASSAPNKNQLDGKKKPKTSPVKPIPQNTTEYTRVRKNADSKNNLN 1800  
DB 1741 FRVKKIMDOVQOASASSAPNKNQLDGKKKPKTSPVKPIPQNTTEYTRVRKNADSKNNLN 1800  
QY 1801 AERVFSDNKSQKQNLKNNSKDFNDKLPNNEDRVRGSAFDPSPHYTPIEGTPYCFSRND 1860  
DB 1801 AERVFSDNKSQKQNLKNNSKDFNDKLPNNEDRVRGSAFDPSPHYTPIEGTPYCFSRND 1860

RESULT 8

AAR63508

ID AAR63508 standard; protein; 2842 AA.

XX

QY 1861 SLSSSLFDDDDDVLSPEKAEIRKAKENKSEBAKVTSHTELTSTNQOSANKTQAIKAPINR 1920  
DB 1861 SLSSSLFDDDDDVLSPEKAEIRKAKENKSEBAKVTSHTELTSTNQOSANKTQAIKAPINR 1920  
QY 1921 GQFKPILQKQSTPPQSSKDIPIRGAATDEKLQNFALIENTPVCFSHNSLSLSDIDQENN 1980  
DB 1921 GQFKPILQKQSTPPQSSKDIPIRGAATDEKLQNFALIENTPVCFSHNSLSLSDIDQENN 1980  
QY 1981 NKENBIDIKETEPDSDGGBSKPOASGYAPKSHVEVTPVCFSRNSLSLSIDSEDDLLQ 2040  
DB 1981 NKENBIDIKETEPDSDGGBSKPOASGYAPKSHVEVTPVCFSRNSLSLSIDSEDDLLQ 2040  
QY 2041 ECTSSAMPKKKPSRLKGDNEKHSPRNMGILGEDJTLDLKDIQRDPDSHGLSPDSENF 2100  
DB 2041 ECTSSAMPKKKPSRLKGDNEKHSPRNMGILGEDJTLDLKDIQRDPDSHGLSPDSENF 2100  
QY 2101 WKAIQEGANSIVSSLHQAAAAACLSRQASDSDSIILSKSGISLGSFPHLTPDQEKPF 2160  
DB 2101 WKAIQEGANSIVSSLHQAAAAACLSRQASDSDSIILSKSGISLGSFPHLTPDQEKPF 2160  
QY 2161 SNKGPRILKPGEKSTLTETTKIESKIGKGGKVKYKSLITGKVRNSSEISGQWKQFLOAN 2220  
DB 2161 SNKGPRILKPGEKSTLTETTKIESKIGKGGKVKYKSLITGKVRNSSEISGQWKQFLOAN 2220  
QY 2221 MPSISGRGTMIHIPGVNRSSTSPVSKGPPPLKTPASKSPSEGTATTSPRGAKPSVK 2280  
DB 2221 MPSISGRGTMIHIPGVNRSSTSPVSKGPPPLKTPASKSPSEGTATTSPRGAKPSVK 2280  
QY 2281 ELSFVARQTSIQGSGKAPSRSGRSDTSPRPAQQLSRPIQSPGNSISPGNGISPPN 2340  
DB 2281 ELSFVARQTSIQGSGKAPSRSGRSDTSPRPAQQLSRPIQSPGNSISPGNGISPPN 2340  
QY 2341 KLSQLPRTSPSTASTKSGSGKMSYTSFGRQVSSQNLTKQTGLSKNASSIPRSEASKG 2400  
DB 2341 KLSQLPRTSPSTASTKSGSGKMSYTSFGRQVSSQNLTKQTGLSKNASSIPRSEASKG 2400  
QY 2401 LNOMNNGANKVKVELSRMSTKSSGSESDRSRPRVLRQSTFIKAPSTLRKLEESA 2460  
DB 2401 LNOMNNGANKVKVELSRMSTKSSGSESDRSRPRVLRQSTFIKAPSTLRKLEESA 2460  
QY 2461 SFESLSPSSRPASPTRSOATPVLSPLDMSLSTHSSVQAGGWRKLPNLSPTIENYNDG 2520  
DB 2461 SFESLSPSSRPASPTRSOATPVLSPLDMSLSTHSSVQAGGWRKLPNLSPTIENYNDG 2520  
QY 2521 RPAKRHDIAARSHSESPRLPINRSGTWKREHSHSSLPVSTWRTTSGSSSILSASSES 2580  
DB 2521 RPAKRHDIAARSHSESPRLPINRSGTWKREHSHSSLPVSTWRTTSGSSSILSASSES 2580  
QY 2581 SEKAKSDEKHVNSISGTQKSKENQVSAKTWKKIKENEFSPNTNSTQTVSSGATNGAES 2640  
DB 2581 SEKAKSDEKHVNSISGTQKSKENQVSAKTWKKIKENEFSPNTNSTQTVSSGATNGAES 2640  
QY 2641 KTLTYQMAPAVSKTEDVWVRIEDCPINNPRSGRSPGTGNTPPVIDSVSEKANPNIKDSKON 2700  
DB 2641 KTLTYQMAPAVSKTEDVWVRIEDCPINNPRSGRSPGTGNTPPVIDSVSEKANPNIKDSKON 2700  
QY 2701 QAKQNVGNGVPMRTVGLNENLNSFTQVADAPDQKGTIEIKPGQNNPVPVSETNESSIVERT 2760  
DB 2701 QAKQNVGNGVPMRTVGLNENLNSFTQVADAPDQKGTIEIKPGQNNPVPVSETNESSIVERT 2760  
QY 2761 PFSSSSSSKHSPPSGTVAARVTPFNPNPSRKSADSTARPQOIPTPVNNNTKKRDSKT 2820  
DB 2761 PFSSSSSSKHSPPSGTVAARVTPFNPNPSRKSADSTARPQOIPTPVNNNTKKRDSKT 2820  
QY 2821 DSTESSGTOQPKRHSGLVTSV 2843  
DB 2821 DSTESSGTOQPKRHSGLVTSV 2843

AC AAR63508;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 23-MAY-1995 (first entry)  
 XX  
 DE Adenomatous polyposis coli tumour repressor.  
 XX  
 KW Adenomatous polyposis coli; tumour repressor; Gardner's syndrome;  
 KW familial adenomatous polyposis; cancer diagnosis and prognosis;  
 KW tumorigenesis suppression.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US5352775-A.  
 XX  
 PD 04-OCT-1994.  
 XX  
 PF 08-AUG-1991; 91US-00741940.  
 XX  
 PR 16-JAN-1991; 91GB-00000962.  
 PR 16-JAN-1991; 91GB-00000963.  
 PR 16-JAN-1991; 91GB-00000974.  
 PR 16-JAN-1991; 91GB-00000975.  
 XX  
 XX (ICIL ) IMPERIAL CHEM IND.  
 PA (CANC-) CANCER INST.  
 PA (UJJO ) UNIV JOHNS HOPKINS.  
 PA (UTAH ) UNIV UTAH.  
 XX  
 PI Markham AF, Hedge PJ, Anand R, Nakamura Y, Groden J, Kinzler K;  
 PI Thliveris A, Carlson M, Vogelstein B, Albertsen H, White RL;  
 PI Joslyn G;  
 XX  
 DR WPI; 1994-316233/39.  
 DR N-PSDB; AAQ72297.  
 XX  
 PT New human adenomatous polyposis coli DNA encoding tumour repressor - and  
 PT derived primers and probes for diagnosis, prognosis and treatment of  
 PT cancer.  
 XX  
 PS Claim 3; Col 71-84; 113pp; English.  
 XX  
 CC AAQ72297 is a cDNA isolated from the human adenomatous polyposis coli  
 CC (APC) gene, it encodes the tumour repressors described in AAR63507 and  
 CC AAR63508. Determination of alterations in APC or its expression products,  
 CC can be used for the diagnosis and prognosis of cancer e.g. colorectal,  
 CC lung and breast tumours; and for determining predisposition to certain  
 CC cancers such as familial adenomatous polyposis (FAP) and Gardner's  
 CC syndrome. The wild type APC gene (or a part of it) can be used  
 CC therapeutically to restore gene function, while primers and probes  
 CC derived from the cDNA (AAQ72333-400 and AAQ72541-568) can be used to  
 CC detect mutations. Also APC proteins or analogues can be administered to  
 CC compensate for a defective gene, and epithelial cells, or transgenic  
 CC animals carrying a mutated APC allele are useful for detecting  
 CC therapeutic agents able to suppress tumorigenesis. (Updated on 25-MAR-  
 CC 2003 to correct PA field.)  
 XX  
 SQ Sequence 2842 AA;  
 Query Match 99.9%; Score 14548.5; DB 2; Length 2842;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2842; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
 QY 1 MAASDYDQLLKQVEALQWENSNLRQLEEDNSNLTTLKLETSANMKVVLKQSGSIEDEAM 60  
 DB 1 MAASDYDQLLKQVEALQWENSNLRQLEEDNSNLTTLKLETSANMKVVLKQSGSIEDEAM 60  
 QY 61 ASSGQIDLLERLKLNDLSDNFPVGVKLRSQWLSRYSVSGREGSVSSRGCSVPVWGSFPR 120  
 DB 61 ASSGQIDLLERLKLNDLSDNFPVGVKLRSQWLSRYSVSGREGSVSSRGCSVPVWGSFPR 120  
 QY 121 RGFVNGSRSTGYLBELEKERSLLADLDKBEKEKWYQAQLNLTFRIDSL-LTENFSL 179

DB 121 RGFVNGSRSTGYLBELEKERSLLADLDKBEKEKWYQAQLNLTFRIDSL-LTENFSL 179  
 QY 181 QDTMTRQLEYEARQIRVAMEEQLGTCQDMKRAORRIARIQIEKIDILRIQLQSQAT 240  
 DB 180 QDTMTRQLEYEARQIRVAMEEQLGTCQDMKRAORRIARIQIEKIDILRIQLQSQAT 239  
 QY 241 EAERSSQNKHETGSHDAERQNEGQVGEINMATSGNGGSGTTRMDHETASVLSSSSTHSA 300  
 DB 240 EAERSSQNKHETGSHDAERQNEGQVGEINMATSGNGGSGTTRMDHETASVLSSSSTHSA 299  
 QY 301 PRRLTSHLGTKEVMYVYLLSMGLTHDKDMSRTLLAMSSSQDSCISMRSQSGCLPLLIQLL 360  
 DB 300 PRRLTSHLGTKEVMYVYLLSMGLTHDKDMSRTLLAMSSSQDSCISMRSQSGCLPLLIQLL 359  
 QY 361 HGNDKDSVLLGNSRGSKEARARASAAALHNIHSPDDKRGRRRIRVHLHLEQIRAYCETC 420  
 DB 360 HGNDKDSVLLGNSRGSKEARARASAAALHNIHSPDDKRGRRRIRVHLHLEQIRAYCETC 419  
 QY 421 WEQEAHEPGMDQDNMPAPVEHQICPACVLMKLSFDEEHRHAMNELGGLQAIABLLQ 480  
 DB 420 WEQEAHEPGMDQDNMPAPVEHQICPACVLMKLSFDEEHRHAMNELGGLQAIABLLQ 479  
 QY 481 VDCMYGLTNDHYSITILRRYAGMALTNLTGCDVANKATLCMKGCMRALVAQLKSESEDL 540  
 DB 480 VDCMYGLTNDHYSITILRRYAGMALTNLTGCDVANKATLCMKGCMRALVAQLKSESEDL 539  
 QY 541 QOVIASVLRLNLSWRADVNSKTLREVGSVKALMECALEKVESTLKSVLASALNLSAHT 600  
 DB 540 QOVIASVLRLNLSWRADVNSKTLREVGSVKALMECALEKVESTLKSVLASALNLSAHT 599  
 QY 601 ENKADICAVDQALAFVLGTLYRSQTNLTALIESGGGILRNVSLSLIATNEDHRQILRENN 660  
 DB 600 ENKADICAVDQALAFVLGTLYRSQTNLTALIESGGGILRNVSLSLIATNEDHRQILRENN 659  
 QY 661 CLOTLLQHLKSHSLTIVSNACGLTNLSARNPKDOEALMDGAVSMKLNLIHSHKHMAM 720  
 DB 660 CLOTLLQHLKSHSLTIVSNACGLTNLSARNPKDOEALMDGAVSMKLNLIHSHKHMAM 719  
 QY 721 GSAAALRNLMANRPAPKVDANIMSPGSSLPSPHVRKOKALEAEIDAQHLSETFDNIDLS 780  
 DB 720 GSAAALRNLMANRPAPKVDANIMSPGSSLPSPHVRKOKALEAEIDAQHLSETFDNIDLS 779  
 QY 781 PKASHRSQKHQSLYGDYVDTNRHDDNDRNDPNTGNTVLPYLNTTVLPSSSSSRGS 840  
 DB 780 PKASHRSQKHQSLYGDYVDTNRHDDNDRNDPNTGNTVLPYLNTTVLPSSSSSRGS 839  
 QY 841 LDSRSSEKDRSLERERGITGLGNVHPATENPCTSSKRGLOISTTAAQIAKWEVEVSAIHTS 900  
 DB 840 LDSRSSEKDRSLERERGITGLGNVHPATENPCTSSKRGLOISTTAAQIAKWEVEVSAIHTS 899  
 QY 901 QEDRSSGTTTELHCVTDERNALRRSSAAHNTSNYNTKTSNRTCSMPYAKLEYKRSS 960  
 DB 900 QEDRSSGTTTELHCVTDERNALRRSSAAHNTSNYNTKTSNRTCSMPYAKLEYKRSS 959  
 QY 961 NDSLNSVSSSDGYGKRGOMKPSIESYSEDDSEKPCSYGOYPADLAHKIHSANHMDNDGE 1020  
 DB 960 NDSLNSVSSSDGYGKRGOMKPSIESYSEDDSEKPCSYGOYPADLAHKIHSANHMDNDGE 1019  
 QY 1021 LDTPIYSLKYSDQLNSGRQSPQNERWARPKHIIIEDEIKQSEQRQSRNQSTTPYVTE 1080  
 DB 1020 LDTPIYSLKYSDQLNSGRQSPQNERWARPKHIIIEDEIKQSEQRQSRNQSTTPYVTE 1079  
 QY 1081 STDDKHLKFPQHFQGOECVSPYRSRGANGSETNRVGSNHGINVNSQLCQEDDYEDDKP 1140  
 DB 1080 STDDKHLKFPQHFQGOECVSPYRSRGANGSETNRVGSNHGINVNSQLCQEDDYEDDKP 1139  
 QY 1141 TNYSERYSSEEEHEEERPTNYSIKYNEEKRVHDQPIDYSLKYATDIPSSQKSPFSKS 1200  
 DB 1140 TNYSERYSSEEEHEEERPTNYSIKYNEEKRVHDQPIDYSLKYATDIPSSQKSPFSKS 1199  
 QY 1201 SSGOSSKTEHMSSESSENTSTPSSNAKQNLHPSACSRSRGOPKATCKVSSINQETIQ 1260  
 DB 1200 SSGOSSKTEHMSSESSENTSTPSSNAKQNLHPSACSRSRGOPKATCKVSSINQETIQ 1259









DB 1800 AERFSDNKKQKLNKNDKFNEDRVRGSAFPHYTPTEGTYCFGRND 1859  
 QY 1861 SLSSLDFFDDVDSLREKAEALRKAKENKESAKYTSHTLTSNOOSANKTOAIKQPINR 1920  
 DB 1860 SLSSLDFFDDVDSLREKAEALRKAKENKESAKYTSHTLTSNOOSANKTOAIKQPINR 1919  
 QY 1921 GQPKILOKQSTFPOSSKXIDPDRGAATDEKLNFAIENTPVCFHNSLSLSIDQENN 1980  
 DB 1920 GQPKILOKQSTFPOSSKXIDPDRGAATDEKLNFAIENTPVCFHNSLSLSIDQENN 1979  
 QY 1981 NKNEPIKETPPDSCGSPSPKASGYAPKPFHVEDTPVCFSRNSSLSSLSIDSEDDLLQ 2040  
 DB 1980 NKNEPIKETPPDSCGSPSPKASGYAPKPFHVEDTPVCFSRNSSLSSLSIDSEDDLLQ 2039  
 QY 2041 ECISAMPKKKPSRLKGDNEKHSRNNGGILGEDTLDLKDIQRPDSEHGLSPDSENF 2100  
 DB 2040 ECISAMPKKKPSRLKGDNEKHSRNNGGILGEDTLDLKDIQRPDSEHGLSPDSENF 2099  
 QY 2101 WKAIQEGANSIVSSLHQAAAACLSQASSDSDSLSLKSGISLSPHLPDQEEKPFT 2160  
 DB 2100 WKAIQEGANSIVSSLHQAAAACLSQASSDSDSLSLKSGISLSPHLPDQEEKPFT 2159  
 QY 2161 SNKGRPRILKPGKSTLETKIESEKGIKGGKVKYKSLITGKVRNSNISISQMKQPLQAN 2220  
 DB 2160 SNKGRPRILKPGKSTLETKIESEKGIKGGKVKYKSLITGKVRNSNISISQMKQPLQAN 2219  
 QY 2221 MFSISGRRTMHIIPGVNRSSSSTSPVSKGPPPLKTPASKSPSEGTATTSRGAKPVS 2280  
 DB 2220 MFSISGRRTMHIIPGVNRSSSSTSPVSKGPPPLKTPASKSPSEGTATTSRGAKPVS 2279  
 QY 2281 ELSPVARTQSLGSGSKAPSRSGSDSTPSRPAQOPLSRPQSPORNISIPRNGISPPN 2340  
 DB 2280 ELSPVARTQSLGSGSKAPSRSGSDSTPSRPAQOPLSRPQSPORNISIPRNGISPPN 2339  
 QY 2341 KLSQLPRTSSPTASTKSGSGKMYTGPGRMSQONLTQKGLSKNASSIPRSESASG 2400  
 DB 2340 KLSQLPRTSSPTASTKSGSGKMYTGPGRMSQONLTQKGLSKNASSIPRSESASG 2399  
 QY 2401 LNQWNGGANKKVLRSMTSKSGSDSDSERPVLVQSFPIKEAPSPILRRKLEESA 2460  
 DB 2400 LNQWNGGANKKVLRSMTSKSGSDSDSERPVLVQSFPIKEAPSPILRRKLEESA 2459  
 QY 2461 SFESLSPSRPASPTRSQAQTPVLSPLPDMSLSTHSSVQAGGWRKLPNLSPTIEYNDG 2520  
 DB 2460 SFESLSPSRPASPTRSQAQTPVLSPLPDMSLSTHSSVQAGGWRKLPNLSPTIEYNDG 2519  
 QY 2521 RPAKHDIARSHSESPRLPINRSGTWKREHSHKSSSLPRVSTWRTGSSSILSASSES 2580  
 DB 2520 RPAKHDIARSHSESPRLPINRSGTWKREHSHKSSSLPRVSTWRTGSSSILSASSES 2579  
 QY 2581 SEKAKSEDEKHVNSISGTQSKENOVSAKGTWRKIKENEFSTNGTSOTVSSGATNGAES 2640  
 DB 2580 SEKAKSEDEKHVNSISGTQSKENOVSAKGTWRKIKENEFSTNGTSOTVSSGATNGAES 2639  
 QY 2641 KTLIYOMAPAVSKTEDVWVRIEDCPINNPRSGRSPGTGTPVIDSVSEKANPNIKDSKON 2700  
 DB 2640 KTLIYOMAPAVSKTEDVWVRIEDCPINNPRSGRSPGTGTPVIDSVSEKANPNIKDSKON 2699  
 QY 2701 QAKQNVGSGVPMRTVGLNRLNSFTOVDAPDKGTETKPGONNPVPVSEINNESSIVERT 2760  
 DB 2700 QAKQNVGSGVPMRTVGLNRLNSFTOVDAPDKGTETKPGONNPVPVSEINNESSIVERT 2759  
 QY 2761 PFSSSSSSKHSPSGTVAARVTPFNPNPSRKSADSTARSQIPTPVNNNTKKRDSKT 2820  
 DB 2760 PFSSSSSSKHSPSGTVAARVTPFNPNPSRKSADSTARSQIPTPVNNNTKKRDSKT 2819  
 QY 2821 DSTESSGTSQPKHSGSYLVTSV 2843  
 DB 2820 DSTESSGTSQPKHSGSYLVTSV 2842

RESULT 10

AAW11922  
 ID AAW11922 standard; protein; 2843 AA.  
 XX  
 AC AAW11922;  
 DT 07-MAY-1997 (first entry)  
 DE Adenomatous polyposis coli protein.  
 XX  
 KW Adenomatous polyposis coli; APC; human; antibody; mutant detection.  
 XX  
 OS Homo sapiens.  
 XX  
 PN JP06347459-A.  
 XX  
 PD 22-DEC-1994.  
 XX  
 PF 07-JUN-1993; 93JP-00136102.  
 XX  
 PR 07-JUN-1993; 93JP-00136102.  
 XX  
 PA (SANK-) SANKO JUNYAKU CO LTD.  
 XX  
 DR WPI; 1995-070957/10.  
 XX  
 PT Detection of abnormal human APC (adenomatous polyposis coli) gene - using antibodies directed against the APC gene product.  
 XX  
 PS Claim 1; Page 7-13; 13pp; Japanese.  
 XX  
 CC This sequence represents the wild type human adenomatous polyposis coli (APC) protein. This sequence, and the APC peptide fragments represented by AAW11923-W11929 can be used in the method of the invention. The method of the invention is for the detection of abnormal APC, using antibodies directed against this sequence, or one of the partial peptide sequences  
 XX  
 SQ Sequence 2843 AA;  
 Query Match 99.9%; Score 14546; DB 2; Length 2843;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 2840; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 MAAASYDQLLKQVEALKWENSNLRQELDENSNLTKLETSANMKVVKLQKQSIIEDEAM 60  
 DB 1 MAAASYDQLLKQVEALKWENSNLRQELDENSNLTKLETSANMKVVKLQKQSIIEDEAM 60  
 QY 61 ASSQCIDLLERLKEINLDSSNFPQVKLRKWSLRSGSREGSVSSRSGECSPPVMSGFPR 120  
 DB 61 ASSQCIDLLERLKEINLDSSNFPQVKLRKWSLRSGSREGSVSSRSGECSPPVMSGFPR 120  
 QY 121 RGFVNGSRESTGYLLELEKERSULLADLDKEEKEKWYQAQNLTKRIDSLPLTENFSL 180  
 DB 121 RGFVNGSRESTGYLLELEKERSULLADLDKEEKEKWYQAQNLTKRIDSLPLTENFSL 180  
 QY 181 QTDMTRRQLEYEARQIRVAMEEQGTCCDMKEKRAQRIARIQIQIEKDILRIQLQSQAT 240  
 DB 181 QTDMTRRQLEYEARQIRVAMEEQGTCCDMKEKRAQRIARIQIQIEKDILRIQLQSQAT 240  
 QY 241 EAERSSQNKHETGSHDAERQNEGQVGEINMATSGNCGQSTTRMDHETASVSSSSTHSA 300  
 DB 241 EAERSSQNKHETGSHDAERQNEGQVGEINMATSGNCGQSTTRMDHETASVSSSSTHSA 300  
 QY 301 PRRLTSHLGTWEMVYSLLSMLGTHDKDDMSRILLAMSSQSDSCISNRQSGCPLLIQLL 360  
 DB 301 PRRLTSHLGTWEMVYSLLSMLGTHDKDDMSRILLAMSSQSDSCISNRQSGCPLLIQLL 360  
 QY 361 HGNDKDSVLLGNRSGSEKARASAAALHNIHSOPDDKRGRRREIRVHLLEQIRAYCETC 420  
 DB 361 HGNDKDSVLLGNRSGSEKARASAAALHNIHSOPDDKRGRRREIRVHLLEQIRAYCETC 420  
 QY 421 WEQEAHEPGMDQKNPMPAPVEHQICFACVCLMKLSFDEEHRHAMNELGGLQIAIELLQ 480  
 DB 421 WEQEAHEPGMDQKNPMPAPVEHQICFACVCLMKLSFDEEHRHAMNELGGLQIAIELLQ 480

QY 481 VDCWYGLTNDHYSLT LRYAGMALTNLTFGDVANKATLCMSKGCMBALVAQLKSESDL 540  
DB 481 VDCWYGLTNDHYSLT LRYAGMALTNLTFGDVANKATLCMSKGCMBALVAQLKSESDL 540  
QY 541 QOVTASVLRLNLSRADVNSKTLREVGSVKALMECALEVKESTLKVLSALWNLNLSAHT 600  
DB 541 QOVTASVLRLNLSRADVNSKTLREVGSVKALMECALEVKESTLKVLSALWNLNLSAHT 600  
QY 601 ENKADI CAVDGALAPLVGTLTYRSQNTWTLAIIESGGGILRVNSSLIATNEDHQILRENN 660  
DB 601 ENKADI CAVDGALAPLVGTLTYRSQNTWTLAIIESGGGILRVNSSLIATNEDHQILRENN 660  
QY 661 CLQTLLOHLKSHSLTIVSNACGTLWNL SARPKDQEBALWDMGAVSMLKNIHSHKWMIAM 720  
DB 661 CLQTLLOHLKSHSLTIVSNACGTLWNL SARPKDQEBALWDMGAVSMLKNIHSHKWMIAM 720  
QY 721 GSAALRLNLMANRPAPKADANIMSPGSLPSLHVKKQKALEAEIDAQHLSETFDNIDNLS 780  
DB 721 GSAALRLNLMANRPAPKADANIMSPGSLPSLHVKKQKALEAEIDAQHLSETFDNIDNLS 780  
QY 781 PKASHRSKQRIKOSLYGDYVDFNTRHDDNRSDNFTGNMTVLPYLNTTVLPSSSSSRGS 840  
DB 781 PKASHRSKQRIKOSLYGDYVDFNTRHDDNRSDNFTGNMTVLPYLNTTVLPSSSSSRGS 840  
QY 841 LDRSRSEKDRSLERERGI GLGNHYHPATENPGTSSKRGLOLSTTAAQIAKYMEEVSAIHTS 900  
DB 841 LDRSRSEKDRSLERERGI GLGNHYHPATENPGTSSKRGLOLSTTAAQIAKYMEEVSAIHTS 900  
QY 901 QEDRSSGSTTELHCVTDERNALRSSAAHSTHNTYNTFKSENSNRCTCSMPYAKLEYKRSS 960  
DB 901 QEDRSSGSTTELHCVTDERNALRSSAAHSTHNTYNTFKSENSNRCTCSMPYAKLEYKRSS 960  
QY 961 NDSLNSVSSSDGYKRGOMKPSIESYEDDESFCYGOVPADLAHKIHSANHMDNDGE 1020  
DB 961 NDSLNSVSSSDGYKRGOMKPSIESYEDDESFCYGOVPADLAHKIHSANHMDNDGE 1020  
QY 1021 LDTPIYNSLYKSDQLNSGRQSPQNRWARPKHIIIEDEIKQSEQRQNRQSTTYPVYTE 1080  
DB 1021 LDTPIYNSLYKSDQLNSGRQSPQNRWARPKHIIIEDEIKQSEQRQNRQSTTYPVYTE 1080  
QY 1081 STDDXHLKLFQHFQOCVCVPSYSRGANGSETNRVGNHGINVQSOLCOEDDYEDDKP 1140  
DB 1081 STDDXHLKLFQHFQOCVCVPSYSRGANGSETNRVGNHGINVQSOLCOEDDYEDDKP 1140  
QY 1141 TNYSERYSBEOHEBERPTNYIKYNEEKRVHDQPIDYSLKYATDIPSSQKQSFSSKS 1200  
DB 1141 TNYSERYSBEOHEBERPTNYIKYNEEKRVHDQPIDYSLKYATDIPSSQKQSFSSKS 1200  
QY 1201 SSGOSSKTEHWSSSENSTPSSNAKQOLHPSSAQSRGQPKAATCKVSSINOETIQ 1260  
DB 1201 SSGOSSKTEHWSSSENSTPSSNAKQOLHPSSAQSRGQPKAATCKVSSINOETIQ 1260  
QY 1261 TYCVEDTPICFSRCSLSLSLSAEDEIGCNQTTQEADSNANTLQIAIEIKKIGTRSAEDPV 1320  
DB 1261 TYCVEDTPICFSRCSLSLSLSAEDEIGCNQTTQEADSNANTLQIAIEIKKIGTRSAEDPV 1320  
QY 1321 SEVPAVQHPRTKSRRLQGSLSLSAESARHKAFFSSGAKSPSKSGAQTPKSPPEHYVQET 1380  
DB 1321 SEVPAVQHPRTKSRRLQGSLSLSAESARHKAFFSSGAKSPSKSGAQTPKSPPEHYVQET 1380  
QY 1381 PLMFSTRCTSVSSLDSPFSRSLIASSVQSEPCSGMVSGIISPDLPSDPTMPSPRSKTPP 1440  
DB 1381 PLMFSTRCTSVSSLDSPFSRSLIASSVQSEPCSGMVSGIISPDLPSDPTMPSPRSKTPP 1440  
QY 1441 PPFQTAQTKREVPNKAPTAKEKSESGKQAAVNAVORVOVLDPADTLHLHFAESTPDGF 1500  
DB 1441 PPFQTAQTKREVPNKAPTAKEKSESGKQAAVNAVORVOVLDPADTLHLHFAESTPDGF 1500  
QY 1501 SCSSLSALSLEDPFFIQKOVELRIMPPVQNDNGNETSEBQPKESNENOEKAEKTI DSE 1560  
DB 1501 SCSSLSALSLEDPFFIQKOVELRIMPPVQNDNGNETSEBQPKESNENOEKAEKTI DSE 1560

QY 1561 KOLLDDSDDDDIIEIEBECIIISAMPTKSSRKAKKPAQTASKLPPPVARKPSOLPVYKLLPS 1620  
DB 1561 KOLLDDSDDDDIIEIEBECIIISAMPTKSSRKAKKPAQTASKLPPPVARKPSOLPVYKLLPS 1620  
QY 1621 QNRLQFQKHVSFTPGDDMPRYCVVEGTPINFSTATSLSDLTIESPPNLAAGEVGRGAQ 1680  
DB 1621 QNRLQFQKHVSFTPGDDMPRYCVVEGTPINFSTATSLSDLTIESPPNLAAGEVGRGAQ 1680  
QY 1681 SGEFEKRDITPTEGRSTDEAOGGKTSSVTIPELDDNKAEBGDIILAEICINSAMPKGSHP 1740  
DB 1681 SGEFEKRDITPTEGRSTDEAOGGKTSSVTIPELDDNKAEBGDIILAEICINSAMPKGSHP 1740  
QY 1741 FRVKIMDOVQOASASSAPKNQOLDGKKKPTSPVKPIPONTYRTRVRKNADSKNIN 1800  
DB 1741 FRVKIMDOVQOASASSAPKNQOLDGKKKPTSPVKPIPONTYRTRVRKNADSKNIN 1800  
QY 1801 AERVFSDNKKQKQNLKXNSKDFNDKLPNNEDVRVGSFAFDSPHHYTPIEGTPTCYFSRND 1860  
DB 1801 AERVFSDNKKQKQNLKXNSKDFNDKLPNNEDVRVGSFAFDSPHHYTPIEGTPTCYFSRND 1860  
QY 1861 SLSLSLDFDDDDVLSREKAEIRKAKENKESAKVTSHTELTSNOOSANKTOAIKAPINR 1920  
DB 1861 SLSLSLDFDDDDVLSREKAEIRKAKENKESAKVTSHTELTSNOOSANKTOAIKAPINR 1920  
QY 1921 GQPKPILOKQSTFPQSSKDIIPDRGAATDEKLQNFPAIENTPVCFSHNSLSLSLSDIDQENN 1980  
DB 1921 GQPKPILOKQSTFPQSSKDIIPDRGAATDEKLQNFPAIENTPVCFSHNSLSLSLSDIDQENN 1980  
QY 1981 NKENEPIKETEPPDQGPSPKQASGVAPKSFHVEDTPVCFSRNSLSLSLSDIDEDLLQ 2040  
DB 1981 NKENEPIKETEPPDQGPSPKQASGVAPKSFHVEDTPVCFSRNSLSLSLSDIDEDLLQ 2040  
QY 2041 ECISAMPKKKPSRLKGDNEKHSPRNMGIGLGEDLTLDKDIORPDSEHGLSPDSENFD 2100  
DB 2041 ECISAMPKKKPSRLKGDNEKHSPRNMGIGLGEDLTLDKDIORPDSEHGLSPDSENFD 2100  
QY 2101 WKAIQEGANSIVSSLHQAAAAACLROASSDSDSLKSGISLGSFPFLTTPDQEEKPFT 2160  
DB 2101 WKAIQEGANSIVSSLHQAAAAACLROASSDSDSLKSGISLGSFPFLTTPDQEEKPFT 2160  
QY 2161 SNKGPRILLKPEKSTLETYKTESKGIKGGKYVYKSLITGKVESNBEISQMKQKQLOAN 2220  
DB 2161 SNKGPRILLKPEKSTLETYKTESKGIKGGKYVYKSLITGKVESNBEISQMKQKQLOAN 2220  
QY 2221 MFSISRGTMIHIFGVNRNSSSTSPVSKGPPKTPASKSPSEGOATTPRGAAPSVKS 2280  
DB 2221 MFSISRGTMIHIFGVNRNSSSTSPVSKGPPKTPASKSPSEGOATTPRGAAPSVKS 2280  
QY 2281 ELSVARTOSIOIGSSKAPSPSGSDSTPSRPAOQPLSRPIQSPGRNSISPGRNGISPPN 2340  
DB 2281 ELSVARTOSIOIGSSKAPSPSGSDSTPSRPAOQPLSRPIQSPGRNSISPGRNGISPPN 2340  
QY 2341 KLSQLPRTSSPSTASTKSSGSKMSYTSFGQMSQQNLTKQTGLSKNASSIPRSESASKG 2400  
DB 2341 KLSQLPRTSSPSTASTKSSGSKMSYTSFGQMSQQNLTKQTGLSKNASSIPRSESASKG 2400  
QY 2401 LNQMNGNANGANKVELSRMSTKSSGSEDSRSEBPVLVRSQTFIKEAPSTFLRKLEESA 2460  
DB 2401 LNQMNGNANGANKVELSRMSTKSSGSEDSRSEBPVLVRSQTFIKEAPSTFLRKLEESA 2460  
QY 2461 SPESLSPSSRPASTRSQAQTFVLSPSLPDMSLSTHSSVQAGGWRKLPPLNLSPTIEYNDG 2520  
DB 2461 SPESLSPSSRPASTRSQAQTFVLSPSLPDMSLSTHSSVQAGGWRKLPPLNLSPTIEYNDG 2520  
QY 2521 RPAKRHDIAIRSHSESPRLPINRSGTWKRBHSHKSSSLPRVSTWRTGTGSSSSILSASSES 2580  
DB 2521 RPAKRHDIAIRSHSESPRLPINRSGTWKRBHSHKSSSLPRVSTWRTGTGSSSSILSASSES 2580  
QY 2581 SEKAKSEDEKHNISIGTKOSKENQVSAKGTWRKIKENEFSPNTSTQTVSSGATNCAES 2640  
DB 2581 SEKAKSEDEKHNISIGTKOSKENQVSAKGTWRKIKENEFSPNTSTQTVSSGATNCAES 2640  
QY 2641 KTLIYQWAPAVSKTDEWVRIEDCPINNRSGRSPTGNTPPVIDSVSEKANPNIKDSKDN 2700

Db 2641 KTLIQMAFAVSKTEDVWRIEDCINPRSGSTGNTFPVIDSVSEKANPNIKDSKN 2700  
QY 2701 QAKQVNGSVPMRTVGLNRLNSFIQVDAPQKTEIKPGQNNPVPVSETNESSIVERT 2760  
Db 2701 QAKQVNGSVPMRTVGLNRLNSFIQVDAPQKTEIKPGQNNPVPVSETNESSIVERT 2760  
QY 2761 PFSSSSSKHSGSPGTVAAARVTPFNPNPSPKSSADSTSARPSQIPTPVNNNTKKRDSKT 2820  
Db 2761 PFSSSSSKHSGSPGTVAAARVTPFNPNPSPKSSADSTSARPSQIPTPVNNNTKKRDSKT 2820  
QY 2821 DTESSSGTQSPKRHSGSYLVTSV 2843  
Db 2821 DTESSSGTQSPKRHSGSYLVTSV 2843

RESULT 11

AD65846  
ID AD65846 standard; protein; 2843 AA.  
XX AC AD65846;  
XX DT 29-JAN-2004 (first entry)  
XX DE Human adenomatosis polyposis coli protein.  
XX KW Human; adenomatosis polyposis coli protein; protein-protein interaction;  
XX KW protein array; PDZ domain; drug target screening.  
XX OS Homo sapiens.

XX US2003170723-A1.  
XX PD 11-SEP-2003.  
XX PF 06-MAR-2002; 2002US-00092138.  
XX PR 06-MAR-2002; 2002US-00092138.  
XX PA (SATO/) SATO T.  
XX PI Sato T;  
XX DR WPI; 2003-852032/79.  
XX PT Preparing a protein array useful for screening drug targets comprises  
PT depositing an array of a first protein on substrate, and applying a  
PT second protein comprising an amino acid sequence that binds to a domain  
PT of the first protein.

XX Disclosure; SEQ ID NO 30; 60pp; English.  
XX CC The invention relates to a method for preparing a protein array based on  
CC protein-protein interaction, by depositing an array of a first protein  
CC comprising a PDZ domain on a substrate, and applying a second protein  
CC comprising an amino acid sequence that binds to the PDZ domain of the  
CC first protein. The method is useful for preparing protein arrays based on  
CC biochemical protein-protein interactions. Arrays produced by this method  
CC are useful for screening drug targets. This sequence represents the human  
CC adenomatosis polyposis coli protein, used in the method of the invention.  
XX Sequence 2843 AA;

Query Watch 99.8%; Score 14539; DB 7; Length 2843;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2837; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
QY 1 MAAASYDQLLKQVEALKMENSRLRQLEDSNHLTKLETEASNKMKVLEVLKQLQGSIEDAM 60  
Db 1 MAAASYDQLLKQVEALKMENSRLRQLEDSNHLTKLETEASNKMKVLEVLKQLQGSIEDAM 60  
QY 61 ASSGQIDLLERLKEINLDSSNPFQVKLRSMKSLRSGSREGSVSSRSGECSPVPMGSPFR 120

Db 61 ASSGQIDLLERLKEINLDSSNPFQVKLRSMKSLRSGSREGSVSSRSGECSPVPMGSPFR 120  
QY 121 RGFVNGSRSTGYLELEKERSILLADLDKEEKEKWYQAOLQNLTKRIDSLPLTENFSL 180  
Db 121 RGFVNGSRSTGYLELEKERSILLADLDKEEKEKWYQAOLQNLTKRIDSLPLTENFSL 180  
QY 181 QTDWTRQLEVEARQIRVAMEBOLGTCQDMKEKAQRRIARIQOIEKDILIRIQLLOSOAT 240  
Db 181 QTDWTRQLEVEARQIRVAMEBOLGTCQDMKEKAQRRIARIQOIEKDILIRIQLLOSOAT 240  
QY 241 EABRSSQNKHETGSHDAERONEGGVGGEINMATSNGQGGSTTRMDHETASVLSSTSTHSA 300  
Db 241 EABRSSQNKHETGSHDAERONEGGVGGEINMATSNGQGGSTTRMDHETASVLSSTSTHSA 300  
QY 301 PRLTSHLGTKEVYVYLLSVLGTGDKDDMSRTLILAMSSQDSCISMRQSGCPLLIQLL 360  
Db 301 PRLTSHLGTKEVYVYLLSVLGTGDKDDMSRTLILAMSSQDSCISMRQSGCPLLIQLL 360  
QY 361 HGNDKDSVLLNGSRGSKAARASAAALHNIHSPQDDKRGRRIRVLLHLEQIRAYCETC 420  
Db 361 HGNDKDSVLLNGSRGSKAARASAAALHNIHSPQDDKRGRRIRVLLHLEQIRAYCETC 420  
QY 421 WEQEAHERGMDQDKNMPAPVEHOICPAVCVLMKLSFDEEHRHAMNELGGLQIAELLQ 480  
Db 421 WEQEAHERGMDQDKNMPAPVEHOICPAVCVLMKLSFDEEHRHAMNELGGLQIAELLQ 480  
QY 481 VDCEMYGLTNDHYSITLRRYAGMALTNLTGFDVANKATLCMKMGCMREALVAQLKSESED 540  
Db 481 VDCEMYGLTNDHYSITLRRYAGMALTNLTGFDVANKATLCMKMGCMREALVAQLKSESED 540  
QY 541 QOVIASVLRNLWRADVNSKTLREVGSVKALMECALEVKEKSTLKSIVLSALWNLSHAHT 600  
Db 541 QOVIASVLRNLWRADVNSKTLREVGSVKALMECALEVKEKSTLKSIVLSALWNLSHAHT 600  
QY 601 ENKADICAVDGAFLVGTLTYSQNTLAIISGGGILNVSLSLATNEDHRIQLRENN 660  
Db 601 ENKADICAVDGAFLVGTLTYSQNTLAIISGGGILNVSLSLATNEDHRIQLRENN 660  
QY 661 CLQTLLOHLKSHSLTIIVSNACGLMNLNARPNKQDEALMDMGAVSMKLIHSHKHVIAM 720  
Db 661 CLQTLLOHLKSHSLTIIVSNACGLMNLNARPNKQDEALMDMGAVSMKLIHSHKHVIAM 720  
QY 721 GSAALENLMANRPKADANIMSPGSSPLSHVRKOKALELDAHLSETFDNDLS 780  
Db 721 GSAALENLMANRPKADANIMSPGSSPLSHVRKOKALELDAHLSETFDNDLS 780  
QY 781 PKASHRSKQHKQSLYGDYVFTNRHDDNRDNFNTGNMTVLPYLNTTTLVLPSSSSSRGS 840  
Db 781 PKASHRSKQHKQSLYGDYVFTNRHDDNRDNFNTGNMTVLPYLNTTTLVLPSSSSSRGS 840  
QY 841 LDSRSRSDRSLEBERGIGLGNYPATENPGTSKRGLOISTTAAQIAKMEVSAIHTS 900  
Db 841 LDSRSRSDRSLEBERGIGLGNYPATENPGTSKRGLOISTTAAQIAKMEVSAIHTS 900  
QY 901 QEDRSSGTTTELHCVTDERNALRRSSAAHTSHNTYNTFTKSENRTCSMPYAKLEYKRSS 960  
Db 901 QEDRSSGTTTELHCVTDERNALRRSSAAHTSHNTYNTFTKSENRTCSMPYAKLEYKRSS 960  
QY 961 NDSLNSVSSSDGYGKRGQMKPSIESYSEDESFCQYGPADLAHKIHSANHMDNDGE 1020  
Db 961 NDSLNSVSSSDGYGKRGQMKPSIESYSEDESFCQYGPADLAHKIHSANHMDNDGE 1020  
QY 1021 LDTPINYSKYSDQNLNSGRQSPQNERWARPKHIIIEDEIKQSEQRQSRNQSTTTPVYTE 1080  
Db 1021 LDTPINYSKYSDQNLNSGRQSPQNERWARPKHIIIEDEIKQSEQRQSRNQSTTTPVYTE 1080  
QY 1081 STDDKHLKFPHFQGCVCVPSYRSGANGSETNRVGSNHGINQVQSQCQEDDYEDDKP 1140  
Db 1081 STDDKHLKFPHFQGCVCVPSYRSGANGSETNRVGSNHGINQVQSQCQEDDYEDDKP 1140  
QY 1141 TNSYRSEBEOHEEERPTNYSIKYNEEKHVDPQIDYSLKYATDIPSSKQSFSEFSK 1200  
Db 1141 TNSYRSEBEOHEEERPTNYSIKYNEEKHVDPQIDYSLKYATDIPSSKQSFSEFSK 1200

QY 1201 SSGOSSKTEHMSSESSENTSTPSSNAKRONLHPSSAQSRSGQPOKAATCKVSSINQETIQ 1260  
 Db 1201 SSGOSSKTEHMSSESSENTSTPSSNAKRONLHPSSAQSRSGQPOKAATCKVSSINQETIQ 1260  
 QY 1261 TYCVEDTPICFSRCSLSSLSAEDIEGICNQTQADSAANTLOIAIEKEKIGTRSAEDPV 1320  
 Db 1261 TYCVEDTPICFSRCSLSSLSAEDIEGICNQTQADSAANTLOIAIEKEKIGTRSAEDPV 1320  
 QY 1321 SEVPAVCHPRTKSRRLQGSLSLSESAHKAHVFSSGAKSPKSGAQTPKSPPEHVVOET 1380  
 Db 1321 SEVPAVCHPRTKSRRLQGSLSLSESAHKAHVFSSGAKSPKSGAQTPKSPPEHVVOET 1380  
 QY 1381 PLMFSRCTSVSSLDSPFSRGIASVQSEPCSGMVSGIISPSDLPDPSGQTMPPSRSKTTP 1440  
 Db 1381 PLMFSRCTSVSSLDSPFSRGIASVQSEPCSGMVSGIISPSDLPDPSGQTMPPSRSKTTP 1440  
 QY 1441 PPPQATQKREVPKNAKAPTAEKRESGPKQAANAARVQVLPDADTLHLHFATESTPDGF 1500  
 Db 1441 PPPQATQKREVPKNAKAPTAEKRESGPKQAANAARVQVLPDADTLHLHFATESTPDGF 1500  
 QY 1501 SCSSLSALSLSDEFFIQKVELRIMPVQENDNGNETESQPKESNENQEKAEKIDSE 1560  
 Db 1501 SCSSLSALSLSDEFFIQKVELRIMPVQENDNGNETESQPKESNENQEKAEKIDSE 1560  
 QY 1561 KOLLDDSDDDIELEBECIISAMPTKSRKAKPAQATASKLPPVAKPSQLPVYKLLPS 1620  
 Db 1561 KOLLDDSDDDIELEBECIISAMPTKSRKAKPAQATASKLPPVAKPSQLPVYKLLPS 1620  
 QY 1621 QNRLOPKHVSFTPGDDMPRVYCVGEGTPIINFSTATSLSDTIESPNELAAGSVRGGAQ 1680  
 Db 1621 QNRLOPKHVSFTPGDDMPRVYCVGEGTPIINFSTATSLSDTIESPNELAAGSVRGGAQ 1680  
 QY 1681 SGEFEKEDTPTTEGRSTDEAOGGKTSSVTIPELDNKAEGDILAEICINSAMPKGSHP 1740  
 Db 1681 SGEFEKEDTPTTEGRSTDEAOGGKTSSVTIPELDNKAEGDILAEICINSAMPKGSHP 1740  
 QY 1741 FRVKKIMDQVOQASASSAPKKNOLDCKKKKPTSPVKPIQNTYEYRVRKKNADSKNNLN 1800  
 Db 1741 FRVKKIMDQVOQASASSAPKKNOLDCKKKKPTSPVKPIQNTYEYRVRKKNADSKNNLN 1800  
 QY 1801 AERFSDNKKONLKNKSKDNFKLNNEDRVGSGFAFDSPHHYTPIEGTYPYCFSRND 1860  
 Db 1801 AERFSDNKKONLKNKSKDNFKLNNEDRVGSGFAFDSPHHYTPIEGTYPYCFSRND 1860  
 QY 1861 SLSSLDFFDDDDVLSREKAEIRKAKENKESFAKVTSHTELTSNOQSANKTQATAKOPINR 1920  
 Db 1861 SLSSLDFFDDDDVLSREKAEIRKAKENKESFAKVTSHTELTSNOQSANKTQATAKOPINR 1920  
 QY 1921 GQPKPILQKSTFPQSSKIDPDRAAATDEKLQNFPAIENTPVCFSHNSLSLSDIDQENN 1980  
 Db 1921 GQPKPILQKSTFPQSSKIDPDRAAATDEKLQNFPAIENTPVCFSHNSLSLSDIDQENN 1980  
 QY 1981 NKNEPIKETEPDPSQSEPKQASGAPKSPHVEDTPVCFSRNSLSLSIDSEDDLQ 2040  
 Db 1981 NKNEPIKETEPDPSQSEPKQASGAPKSPHVEDTPVCFSRNSLSLSIDSEDDLQ 2040  
 QY 2041 ECISAMPKPKKPRKLGDNKXHPNMGILGEBDLTLDKIDQRPDSEHGLSPDSENF 2100  
 Db 2041 ECISAMPKPKKPRKLGDNKXHPNMGILGEBDLTLDKIDQRPDSEHGLSPDSENF 2100  
 QY 2101 WKATQEGANSIVSSLHQAACALSRQASSDSLSLSKSGISLGSFHLTPDQEEKPFT 2160  
 Db 2101 WKATQEGANSIVSSLHQAACALSRQASSDSLSLSKSGISLGSFHLTPDQEEKPFT 2160  
 QY 2161 SNKGPRILKPEKSTLTETKTESKGIKGGKVKYKSLITGKVRNSEISGQWKQLOAN 2220  
 Db 2161 SNKGPRILKPEKSTLTETKTESKGIKGGKVKYKSLITGKVRNSEISGQWKQLOAN 2220  
 QY 2221 MFSISRGRMTIHIPVNRNSSSSTSPVSKGPPPLKTPASKSPSEGTATTSPRGAKPSVKS 2280  
 Db 2221 MFSISRGRMTIHIPVNRNSSSSTSPVSKGPPPLKTPASKSPSEGTATTSPRGAKPSVKS 2280

QY 2281 ELSPVARTSOIGSSSKAPSRSGSRDSTPSPAPQPLSRPIQSPGRNSISPGRNGISPPN 2340  
 Db 2281 ELSPVARTSOIGSSSKAPSRSGSRDSTPSPAPQPLSRPIQSPGRNSISPGRNGISPPN 2340  
 QY 2341 KLSQLPRTSSSTASTKSSGSKMSYTPGQMSQONLTQOTGLSKNASSITPSESASKG 2400  
 Db 2341 KLSQLPRTSSSTASTKSSGSKMSYTPGQMSQONLTQOTGLSKNASSITPSESASKG 2400  
 QY 2401 LQNMNNGANKKVELSEMSSTKSSGSDSRSEPRVLVRQSTFIKEAPSPTLRKLBEA 2460  
 Db 2401 LQNMNNGANKKVELSEMSSTKSSGSDSRSEPRVLVRQSTFIKEAPSPTLRKLBEA 2460  
 QY 2461 SFESLSPSRPASPTRSOAQTIVLSPLPMSLTHSSVQAGGWRKLPNLSPTIEYNDG 2520  
 Db 2461 SFESLSPSRPASPTRSOAQTIVLSPLPMSLTHSSVQAGGWRKLPNLSPTIEYNDG 2520  
 QY 2521 RPAKCHDITARGHSESPSRPLPINRSGTWKREHSHKSSSLPRVSTWRTGSSSSILSASES 2580  
 Db 2521 RPAKCHDITARGHSESPSRPLPINRSGTWKREHSHKSSSLPRVSTWRTGSSSSILSASES 2580  
 QY 2581 SEKAKSEDEKHVNSISGKQKQENOVAKGTWRKI KENEFSPNTNSTQTVSSGATNGAES 2640  
 Db 2581 SEKAKSEDEKHVNSISGKQKQENOVAKGTWRKI KENEFSPNTNSTQTVSSGATNGAES 2640  
 QY 2641 KTLIYQMAPAVSKTEDVWRIEDCPINNPRSGRPTGNTPPVIDSVSEKANPNIKDSKDN 2700  
 Db 2641 KTLIYQMAPAVSKTEDVWRIEDCPINNPRSGRPTGNTPPVIDSVSEKANPNIKDSKDN 2700  
 QY 2701 QAKQVNGSGVPMRTVGLNRLNSFIQVADAPDQKGTIKPGQNNPVPVSETNESSIVERT 2760  
 Db 2701 QAKQVNGSGVPMRTVGLNRLNSFIQVADAPDQKGTIKPGQNNPVPVSETNESSIVERT 2760  
 QY 2761 PFSSSSSKHSPSGTVAARVTPEYNPSPEKSADSTSAAPSOIPTPVNNNTKKRDKT 2820  
 Db 2761 PFSSSSSKHSPSGTVAARVTPEYNPSPEKSADSTSAAPSOIPTPVNNNTKKRDKT 2820  
 QY 2821 DSTESSGTQSPKSHSGSYLTVSV 2843  
 Db 2821 DSTESSGTQSPKSHSGSYLTVSV 2843  
 RESULT 12  
 AAR26052  
 ID AAR26052 standard; protein; 2843 AA.  
 AC AAR26052;  
 AC AAR26052;  
 DT 25-MAR-2003 (revised)  
 DT 28-JAN-1993 (first entry)  
 XX  
 XX APC gene product in familial adenomatous polyposis.  
 KW neoplasm; cancer; oncogene; tumour; growth; detection; diagnosis;  
 KW prognosis; treatment; sporadic colorectal carcinomas; ss.  
 OS Homo sapiens.  
 XX  
 XX WO9213103-A1.  
 XX  
 XX 06-AUG-1992.  
 XX  
 XX 16-JAN-1992; 92WO-US000376.  
 XX  
 XX 16-JAN-1991; 91CB-00000963.  
 PR 08-AUG-1991; 91US-00741940.  
 XX  
 XX (UKJO ) UNIV JOHNS HOPKINS.  
 PA (ICIL ) IMPERIAL CHEM IND PLC.  
 PA (UTAH ) UNIV UTAH.  
 PA (CANC-) CANCER INST.  
 XX  
 XX Kinzler KW, Vogelstein B, Anand R, Hedge PJ, Markham AF;  
 PI Albertsen H, Carlson ML, Groden JL, Joslyn G, Thliveris A, White RL;

PI Nakamura Y;  
XX WPI; 1992-284685/34.  
DR N-PSDB; AAQ27234.  
XX  
XX Detection of somatic and germ-line alterations of human APC gene - used  
PT to diagnose, treat and study familial adenomatous polyposis and sporadic  
PT colorectal cancer.  
XX  
XX Disclosure; Page 47; 132pp; English.  
XX  
XX This sequence is encoded by the APC (Adenomatous Polyposis Coli) gene  
CC associated with tumorigenesis, found on chromosome 5q. The sequence may  
CC be mutated by deletions, insertions, inversions, or point mutations of the  
CC gene. The APC gene is expressed in most normal tissues as well suggesting  
CC that APC is a tumour suppressor. (Updated on 25-MAR-2003 to correct PN  
CC field.) (Updated on 25-MAR-2003 to correct PI field.)  
XX  
XX Sequence 2843 AA;

Query Match 99.8%; Score 14533; DB 2; Length 2843;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2836; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
  
QY 1 MAAASYDQLLKQVEALKMENSRLROELEDNSNHLTKLETSANMKVEVLKQLQGSIEDEAM 60  
DB 1 MAAASYDQLLKQVEALKMENSRLROELEDNSNHLTKLETSANMKVEVLKQLQGSIEDEAM 60  
  
QY 61 ASSGOIDLRLERKELNLDSSNPPGKLRKMSLRYSRGSSVSRSGECSPVPMGSPPR 120  
DB 61 ASSGOIDLRLERKELNLDSSNPPGKLRKMSLRYSRGSSVSRSGECSPVPMGSPPR 120  
  
QY 121 RGFVNGSRSTGYLBELEKERSILLADLDKEKEDWYQAOLNLTKRIDSPLTFENFSL 180  
DB 121 RGFVNGSRSTGYLBELEKERSILLADLDKEKEDWYQAOLNLTKRIDSPLTFENFSL 180  
  
QY 181 QTDLTRRQLEYEARQIRVAMEEQLGTCODMEXRAQRRIARIQIEKDLIRLQLOSQAT 240  
DB 181 QTDLTRRQLEYEARQIRVAMEEQLGTCODMEXRAQRRIARIQIEKDLIRLQLOSQAT 240  
  
QY 241 EAERSQKHETGSHDAERQNEGGVGEINMATSGGQSTTRMDHETASVLSSTSTHSA 300  
DB 241 EAERSQKHETGSHDAERQNEGGVGEINMATSGGQSTTRMDHETASVLSSTSTHSA 300  
  
QY 301 PRLTSLHGTQKVMYVLSLMLGTHDKDDMSRTLLAMSSQDSCISMRQSGCLPLLIQLL 360  
DB 301 PRLTSLHGTQKVMYVLSLMLGTHDKDDMSRTLLAMSSQDSCISMRQSGCLPLLIQLL 360  
  
QY 361 HGNDKDSVLLGNSRGSKEARAPASAAALHNIHSQPDCKGRREIRVLHLBOIRAYCETC 420  
DB 361 HGNDKDSVLLGNSRGSKEARAPASAAALHNIHSQPDCKGRREIRVLHLBOIRAYCETC 420  
  
QY 421 WZQEAHEPGMDQDNMPAPVEHQICPAVCVLMKLSFDEEHRHAMNELGGIQAIAELLQ 480  
DB 421 WZQEAHEPGMDQDNMPAPVEHQICPAVCVLMKLSFDEEHRHAMNELGGIQAIAELLQ 480  
  
QY 481 VDCMYGLTNDHYSITLRYAGNALTNLTFGDVANKATILCSMKGCMRALVAQLKSESDL 540  
DB 481 VDCMYGLTNDHYSITLRYAGNALTNLTFGDVANKATILCSMKGCMRALVAQLKSESDL 540  
  
QY 541 QOVIASVLNLSWRADVNSKTLREVGSVKALMECALEVKKESTILKSVLSALWNLSAHCT 600  
DB 541 QOVIASVLNLSWRADVNSKTLREVGSVKALMECALEVKKESTILKSVLSALWNLSAHCT 600  
  
QY 601 ENKADI CAVDGALAFVLGTLTYRSQNTTLAIIESGGILRVNYSLLIATNEDHRQLRENN 660  
DB 601 ENKADI CAVDGALAFVLGTLTYRSQNTTLAIIESGGILRVNYSLLIATNEDHRQLRENN 660  
  
QY 661 CLQTLLOHLKSHSLTIVSNACQGLTNLSARNPKDQBALMDGAVSMLKNLILSHKHMIA 720  
DB 661 CLQTLLOHLKSHSLTIVSNACQGLTNLSARNPKDQBALMDGAVSMLKNLILSHKHMIA 720  
  
QY 721 GSAALRLNLMANRPARYKQKALEAELDAQHLSETFDNDLS 780

DB 721 GSAALRLNLMANRPARYKQKALEAELDAQHLSETFDNDLS 780  
QY 781 PKASHRSKQKHOSLQVGYVDFNRHDDNRSDNFNTGNMTVLSPYLNTTTLVLPSSSSSRGS 840  
DB 781 PKASHRSKQKHOSLQVGYVDFNRHDDNRSDNFNTGNMTVLSPYLNTTTLVLPSSSSSRGS 840  
QY 841 LDSRSSEKRSLEREERGIGLGNYPATENPGTSSKGLQISTTAAQIAKVMBSVAIHTS 900  
DB 841 LDSRSSEKRSLEREERGIGLGNYPATENPGTSSKGLQISTTAAQIAKVMBSVAIHTS 900  
QY 901 QEDRSSSTTELHCVTDERNALRRSSAAHSTNYNFTKSENENRRCSCMPYAKLEYKRSS 960  
DB 901 QEDRSSSTTELHCVTDERNALRRSSAAHSTNYNFTKSENENRRCSCMPYAKLEYKRSS 960  
QY 961 NDSLSNVSSSDGYGKRGQMKPSIESYSEDESDESKFCYGYQVPADLAHKIHSANMDNDGE 1020  
DB 961 NDSLSNVSSSDGYGKRGQMKPSIESYSEDESDESKFCYGYQVPADLAHKIHSANMDNDGE 1020  
QY 1021 LDPINYSKYSDOLNSGRQSPQNERWARPKHIIEDIKQSFQSRNQSTTYPYTE 1080  
DB 1021 LDPINYSKYSDOLNSGRQSPQNERWARPKHIIEDIKQSFQSRNQSTTYPYTE 1080  
QY 1081 STDDKHLKFPQHFQGCQECVSPYSRGANGSETNRVGSNHGNGINQVQSLSQLEDDEYDKP 1140  
DB 1081 STDDKHLKFPQHFQGCQECVSPYSRGANGSETNRVGSNHGNGINQVQSLSQLEDDEYDKP 1140  
QY 1141 TNSERYSEBQHEEERPTNYSIKYNEEKRHVDQIDYSLKYATDIPSSQKQSFSPSKS 1200  
DB 1141 TNSERYSEBQHEEERPTNYSIKYNEEKRHVDQIDYSLKYATDIPSSQKQSFSPSKS 1200  
QY 1201 SSGQSKTEHWSSENSTSTPSSNAKQNLHPSQAQSRGQPKAATCKVSSINQETIQ 1260  
DB 1201 SSGQSKTEHWSSENSTSTPSSNAKQNLHPSQAQSRGQPKAATCKVSSINQETIQ 1260  
QY 1261 TYCVEDTPICFRCSLSLSLSSAEDIGCQNTTQADASANTLQIAIEKIKIGTRSAEDPV 1320  
DB 1261 TYCVEDTPICFRCSLSLSLSSAEDIGCQNTTQADASANTLQIAIEKIKIGTRSAEDPV 1320  
QY 1321 SEVPAVSCHERTKSSRLQGSLSLSSAARHKAFFSSGAKSPSKGQATPKSPHYVQET 1380  
DB 1321 SEVPAVSCHERTKSSRLQGSLSLSSAARHKAFFSSGAKSPSKGQATPKSPHYVQET 1380  
QY 1381 PLMFRCSTSVSSLDSPESRSIASSVQSEPCSGMVSGIISPSDLPSPGQTMPPSRSKTTP 1440  
DB 1381 PLMFRCSTSVSSLDSPESRSIASSVQSEPCSGMVSGIISPSDLPSPGQTMPPSRSKTTP 1440  
QY 1441 PPQTAQTKREVKNKAPTAKEESGPKQAANAAVQVRQVLPDADTLHLFAFESTPDGF 1500  
DB 1441 PPQTAQTKREVKNKAPTAKEESGPKQAANAAVQVRQVLPDADTLHLFAFESTPDGF 1500  
QY 1501 SCSSLSALSILDEPFIQKDVLELIMPPVQENDNGNETESQPKESNENQKEAEKIDSE 1560  
DB 1501 SCSSLSALSILDEPFIQKDVLELIMPPVQENDNGNETESQPKESNENQKEAEKIDSE 1560  
QY 1561 KDLDDSDDDDIIELEECIIISAMPTKSSRKAKKPAQATASKLPPPPVARKPSQLPVYKLLPS 1620  
DB 1561 KDLDDSDDDDIIELEECIIISAMPTKSSRKAKKPAQATASKLPPPPVARKPSQLPVYKLLPS 1620  
QY 1621 QNRLOPKQHVSTPGDDMPRVYCVGTPINFSTATSLDLTIESPPNELAAGVGVRGAQ 1680  
DB 1621 QNRLOPKQHVSTPGDDMPRVYCVGTPINFSTATSLDLTIESPPNELAAGVGVRGAQ 1680  
QY 1681 SGPEFKRDTIPTTEGRSTDEAQQGKTSSVTIPELDDNKAEEGDILASCINSAMPKGSKHP 1740  
DB 1681 SGPEFKRDTIPTTEGRSTDEAQQGKTSSVTIPELDDNKAEEGDILASCINSAMPKGSKHP 1740  
QY 1741 FRVKKINDVQQAASASSAPKNQLDGKKKPTSPYKPIQNTTEYTRVRKNADSKNNLN 1800  
DB 1741 FRVKKINDVQQAASASSAPKNQLDGKKKPTSPYKPIQNTTEYTRVRKNADSKNNLN 1800  
QY 1801 AERFSDNCKSKONLKNNSKDFNDKLPNNEDRVGSAFDSPPHYTPIEGTYPYCFSRND 1860

Db 1801 AERFVSNKDKSKKLNKNNKDNKNDKLPNNEDVRGSPAFDSPHHYTPIEGTFCFSRND 1860  
 Qy 1861 SLSLDDDDVLSRKAELRKAENKESEAKVTSHTELTNSQOQANKTQAIKQPINR 1920  
 Db 1861 SLSLDDDDVLSRKAELRKAENKESEAKVTSHTELTNSQOQANKTQAIKQPINR 1920  
 Qy 1921 GQPKPILOKOSTPQSSKDIPIRGAATDEKQNFATENTPVCFSHNSLSLSIDIDQENN 1980  
 Db 1921 GQPKPILOKOSTPQSSKDIPIRGAATDEKQNFATENTPVCFSHNSLSLSIDIDQENN 1980  
 Qy 1981 NKNEPIKETEPDQSGEPKQASGVAPKSFHVEPTPVCFSRNSLSLSIDSEDDLIQ 2040  
 Db 1981 NKNEPIKETEPDQSGEPKQASGVAPKSFHVEPTPVCFSRNSLSLSIDSEDDLIQ 2040  
 Qy 2041 ECISAMPKKKPSRLKGDNEKSPRNWGGILGEDLTLDLKDIOQRPDSEHGLSPDSENF 2100  
 Db 2041 ECISAMPKKKPSRLKGDNEKSPRNWGGILGEDLTLDLKDIOQRPDSEHGLSPDSENF 2100  
 Qy 2101 WKAIQEGANSIVSSLHQAACLSRQASDSDSILSLKSGISLGSPFHLTPDQEKPT 2160  
 Db 2101 WKAIQEGANSIVSSLHQAACLSRQASDSDSILSLKSGISLGSPFHLTPDQEKPT 2160  
 Qy 2161 SNKGPRILKGEKSTLTETKIEBESKIKGGKVKYKSLITGKVRNSSETSGQWKQLOAN 2220  
 Db 2161 SNKGPRILKGEKSTLTETKIEBESKIKGGKVKYKSLITGKVRNSSETSGQWKQLOAN 2220  
 Qy 2221 MPSISRGRTHIIPGVNNSSTSPVSKGPPPLKTPASKSPSEGQATTSPRGAKEVKS 2280  
 Db 2221 MPSISRGRTHIIPGVNNSSTSPVSKGPPPLKTPASKSPSEGQATTSPRGAKEVKS 2280  
 Qy 2281 ELSVARTQTOIGSSKAPRSGSRDTPSRPAQOPLSRPIQSPGRNSIPGRNGISPPN 2340  
 Db 2281 ELSVARTQTOIGSSKAPRSGSRDTPSRPAQOPLSRPIQSPGRNSIPGRNGISPPN 2340  
 Qy 2341 KLSQLPRTSPSTASTKSSGKMSYTSRQWQSQNLTKQTGLSKNASSIPRSEASKG 2400  
 Db 2341 KLSQLPRTSPSTASTKSSGKMSYTSRQWQSQNLTKQTGLSKNASSIPRSEASKG 2400  
 Qy 2401 LNMWNGANGANKVELSRMSTKSSGESDRSERPVLVROSTFIKAPSPTLRKLKESA 2460  
 Db 2401 LNMWNGANGANKVELSRMSTKSSGESDRSERPVLVROSTFIKAPSPTLRKLKESA 2460  
 Qy 2461 SFESLSPSSRPASPTRQAOTPVLSPLDMSLSTHSSVOAGGWRKLPNLSPTIENDG 2520  
 Db 2461 SFESLSPSSRPASPTRQAOTPVLSPLDMSLSTHSSVOAGGWRKLPNLSPTIENDG 2520  
 Qy 2521 RPAKRHDIAHSHSPSRPLPINSRGTWKRSHKSSSLPRVSTWRTGTSSSILSASSES 2580  
 Db 2521 RPAKRHDIAHSHSPSRPLPINSRGTWKRSHKSSSLPRVSTWRTGTSSSILSASSES 2580  
 Qy 2581 SEKAKSEDEKHVNSISGTQSKENQVSAKGTWKIKENEPSTNSQTSVSSGATNGAES 2640  
 Db 2581 SEKAKSEDEKHVNSISGTQSKENQVSAKGTWKIKENEPSTNSQTSVSSGATNGAES 2640  
 Qy 2641 KTLIYQAPAVSKTDEYVWRIEDCPINPRSGRSPGTNTPPVIDSVSEKANPNIKDKDN 2700  
 Db 2641 KTLIYQAPAVSKTDEYVWRIEDCPINPRSGRSPGTNTPPVIDSVSEKANPNIKDKDN 2700  
 Qy 2701 QAKQNVGNGVPMRTVGLNRLNSFIQVADPQKGTIEKPGQNNPVPVSTNETSSIVERT 2760  
 Db 2701 QAKQNVGNGVPMRTVGLNRLNSFIQVADPQKGTIEKPGQNNPVPVSTNETSSIVERT 2760  
 Qy 2761 PFSSSSSKKSSPSGCTVAARVTPPNYNPSRKSADSTSPRQIPTPVNNNTKDDSKT 2820  
 Db 2761 PFSSSSSKKSSPSGCTVAARVTPPNYNPSRKSADSTSPRQIPTPVNNNTKDDSKT 2820  
 Qy 2821 DSTESSGTQSPKRHSGSYLTVS 2843  
 Db 2821 DSTESSGTQSPKRHSGSYLTVS 2843

ID AAW35392 standard; protein; 2843 AA.  
 XX AAW35392;  
 XX AC  
 XX 25-MAR-2003 (revised)  
 DT 11-MAR-1998 (first entry)  
 DT  
 XX Human adenomatous Polyposis coli gene product.  
 DE  
 XX Human; adenomatous Polyposis coli; APC; diagnosis; prognosis;  
 KW neoplastic tissue; tumour tissue; tumour repressor; mutation;  
 KW sporadic colorectal cancer; detection.  
 XX  
 OS Homo sapiens.  
 XX  
 XX US5648212-A.  
 PN  
 XX 15-JUL-1997.  
 PD  
 XX 12-AUG-1994; 94US-00289548.  
 XX  
 PR 16-JAN-1991; 91GB-00000962.  
 PR 16-JAN-1991; 91GB-00000963.  
 PR 16-JAN-1991; 91GB-00000974.  
 PR 16-JAN-1991; 91GB-00000975.  
 PR 08-AUG-1991; 91US-00741940.  
 XX  
 XX (NICA-) JAPANESE FOUND CANCER RES.  
 PA  
 PA (UTAH) UNIV UTAH.  
 PA (UIJO) UNIV JOHNS HOPKINS.  
 PA (ZENE) ZENECA LTD.  
 XX  
 XX Markham A, Nakamura Y, Groden J, Carlson M, Kinzler K;  
 PI Albertsen H, Hedge PJ, Vogelstein B, Thliveris A, Anand R, White RL;  
 PI Joslyn G;  
 XX  
 XX WPI: 1997-372053/34.  
 DR N-PSDB; AAR95538.  
 XX  
 XX Cancer diagnosis - by detecting mutation(s) in adenomatous polyposis coli  
 PT gene.  
 XX  
 XX Example 1; Col 33-52; 140pp; English.  
 PS  
 XX The present sequence is the human adenomatous Polyposis coli (APC) gene  
 CC product, which was used in the development of a novel method of  
 CC diagnosing or prognosing an APC gene associated neoplastic tissue. The  
 CC method comprises comparing APC gene coding sequences or mRNA in a tumour  
 CC tissue, to APC gene coding sequences or mRNA in a non-neoplastic tissue,  
 CC where a difference indicates an APC gene associated neoplasia of the  
 CC tumour tissue. APC is a tumour repressor expressed in most normal  
 CC tissues. APC mutations are found in familial adenomatous polyposis and  
 CC sporadic colorectal cancer patients. The method enables mutations to be  
 CC detected to provide an indication of predisposition to cancer. (Updated  
 CC on 25-MAR-2003 to correct PR field.)  
 XX  
 XX Sequence 2843 AA;

Query Match 99.8%; Score 14533; DB 2; Length 2843;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 2836; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
 Qy 1 MAASDYDQLLKQVEALKMENSNLQAELEDNSHLTKLETEASNKKEVLKQIQGSIEDEAM 60  
 Db 1 MAASDYDQLLKQVEALKMENSNLQAELEDNSHLTKLETEASNKKEVLKQIQGSIEDEAM 60  
 Qy 61 ASSGQIDLLERLKEINLDSSNFPQVKLRKMSLSYSGRSGSVSRGECSPVPMGSPFR 120  
 Db 61 ASSGQIDLLERLKEINLDSSNFPQVKLRKMSLSYSGRSGSVSRGECSPVPMGSPFR 120  
 Qy 121 RGFVNGSRSTGYLEELERKERSLLIADLDKEEKDWYAOQLNLTKRIDSPLTENFSL 180  
 Db 121 RGFVNGSRSTGYLEELERKERSLLIADLDKEEKDWYAOQLNLTKRIDSPLTENFSL 180



181 QTMTRRQLEYEARQIRVAMEEQLGTCODMEKRAQRIARI:QOIEKOILRIRQLLOQAT 240  
 181 QTDLTRRQLEYEARQIRVAMEEQLGTCODMEKRAQRIARI:QOIEKOILRIRQLLOQAT 240  
 241 EAERSQNHETGSHDARONEGQGVGINNATSGOGGSTRMDHETASVLSSTHSA 300  
 241 EAERSQNHETGSHDARONEGQGVGINNATSGOGGSTRMDHETASVLSSTHSA 300  
 301 PERLTSHLGTKVEMYVYLLSMLGTHDKDDMSRTLLAMSSQDSCISMFQSCCLPILLIQLL 360  
 301 PERLTSHLGTKVEMYVYLLSMLGTHDKDDMSRTLLAMSSQDSCISMFQSCCLPILLIQLL 360  
 361 HGNDKDSVLLGNSRGSKAPARASAAALHNI:HSQDDXGRREIRVLHLLQIRAYCETC 420  
 361 HGNDKDSVLLGNSRGSKAPARASAAALHNI:HSQDDXGRREIRVLHLLQIRAYCETC 420  
 421 WENQEAHEPFGDODKNPMPAPVEHQICPAVCVLMKLSFDEBHRHAMNELGLQIAELQ 480  
 421 WENQEAHEPFGDODKNPMPAPVEHQICPAVCVLMKLSFDEBHRHAMNELGLQIAELQ 480  
 481 VDCENYGLTNDHYSITLRYAGMALTNITFGDVANKATLCNSKGMALVAQLKSESD 540  
 481 VDCENYGLTNDHYSITLRYAGMALTNITFGDVANKATLCNSKGMALVAQLKSESD 540  
 541 QOVIASVLNLSWRADVNSKTLREVGSVKALMECALEVKKESTLKSVLKALWNLASHCT 600  
 541 QOVIASVLNLSWRADVNSKTLREVGSVKALMECALEVKKESTLKSVLKALWNLASHCT 600  
 601 ENKADICAVDGAFLVGLTLYRSQTNLTALIESGGILRNVSLSIATNEDHROILRENN 660  
 601 ENKADICAVDGAFLVGLTLYRSQTNLTALIESGGILRNVSLSIATNEDHROILRENN 660  
 661 CLQTLLOHLKSHSLTIVSNACGTLNLSARNPKDOEALWDMGAVSMKLNLI:HSKXMTAM 720  
 661 CLQTLLOHLKSHSLTIVSNACGTLNLSARNPKDOEALWDMGAVSMKLNLI:HSKXMTAM 720  
 721 GSAALRLNMANRPKAKYKADANIMSGSSLPSSLVHVKQKALAEALDAQHLSEITFNIDNLS 780  
 721 GSAALRLNMANRPKAKYKADANIMSGSSLPSSLVHVKQKALAEALDAQHLSEITFNIDNLS 780  
 781 PKASHRSKORHQSLYGDYVFTNDRHNDNRDNFTGNMTVLSPLYNTTVLPSSSSSRGS 840  
 781 PKASHRSKORHQSLYGDYVFTNDRHNDNRDNFTGNMTVLSPLYNTTVLPSSSSSRGS 840  
 841 LQSSRSKDRSLERERGLGNYPHATENPGTSSKRGLOISTTAAQIAKWEEVSAIHTS 900  
 841 LQSSRSKDRSLERERGLGNYPHATENPGTSSKRGLOISTTAAQIAKWEEVSAIHTS 900  
 901 QEDRSSTTELHCVTDERNALRRSSAAHTHNTYNTFTKSENSNRTCSEMPYAKLEYKRSS 960  
 901 QEDRSSTTELHCVTDERNALRRSSAAHTHNTYNTFTKSENSNRTCSEMPYAKLEYKRSS 960  
 961 NQSLNSVSSDGYKQKQKPSIESYDEDESKFCSYQYPADLAHKTHSANHMDNDGE 1020  
 961 NQSLNSVSSDGYKQKQKPSIESYDEDESKFCSYQYPADLAHKTHSANHMDNDGE 1020  
 1021 LQTPINYSILKYDEQLNSGRQSPQNERWARPKHIEDEIKQSEQRQSRNQSTTYPVYTE 1080  
 1021 LQTPINYSILKYDEQLNSGRQSPQNERWARPKHIEDEIKQSEQRQSRNQSTTYPVYTE 1080  
 1081 STDDKHLKQPHFGQOECPVPSRGRGANGSETNRVGSNHGINQNSQSLCOEDDIEDDKP 1140  
 1081 STDDKHLKQPHFGQOECPVPSRGRGANGSETNRVGSNHGINQNSQSLCOEDDIEDDKP 1140  
 1141 TNYSERYSBEEQHEERPTNYSIKYNEEKRVHVDOPIDYSLKYATDIPSSOKQSPFSKS 1200  
 1141 TNYSERYSBEEQHEERPTNYSIKYNEEKRVHVDOPIDYSLKYATDIPSSOKQSPFSKS 1200  
 1201 SSGQSSKTEHMSSESSNTSTPSSNAKRONQLHPSSAQSRSGQPQAAATCKVSSINQETIQ 1260  
 1201 SSGQSSKTEHMSSESSNTSTPSSNAKRONQLHPSSAQSRSGQPQAAATCKVSSINQETIQ 1260

1261 TYCVETPTICFSSCSLSLSAEDBIGNOTTQEADSANTLOIAIBIKKIGTRSAEDPV 1320  
 1261 TYCVETPTICFSSCSLSLSAEDBIGNOTTQEADSANTLOIAIBIKKIGTRSAEDPV 1320  
 1321 SEVPAYSQBPRTKSRLOQSSLSSESARHKAVEFSSGAKSPSKSGAQTPKSPPEHYVQET 1380  
 1321 SEVPAYSQBPRTKSRLOQSSLSSESARHKAVEFSSGAKSPSKSGAQTPKSPPEHYVQET 1380  
 1381 PLMFSTRCTSVSSLDSPESRSIASSVSQPCSGMVSGIISPSDLDPSPGOTMPPSRSKTTP 1440  
 1381 PLMFSTRCTSVSSLDSPESRSIASSVSQPCSGMVSGIISPSDLDPSPGOTMPPSRSKTTP 1440  
 1441 PPQTAQTKKEVPKNKAPTAEKESGPKQAAVNAVQVQVLPDADTLHFAFATESPDGF 1500  
 1441 PPQTAQTKKEVPKNKAPTAEKESGPKQAAVNAVQVQVLPDADTLHFAFATESPDGF 1500  
 1501 SCSSLSALSDELFFIQQOVLELIMPPVQENDNGNETESQPKESNENQKEAEKTIDSE 1560  
 1501 SCSSLSALSDELFFIQQOVLELIMPPVQENDNGNETESQPKESNENQKEAEKTIDSE 1560  
 1561 KDLLDSDDDDDIIEBECIISAMPTKSSRKAKKPAQATASKLPPPVARKSOLPVYKLLPS 1620  
 1561 KDLLDSDDDDDIIEBECIISAMPTKSSRKAKKPAQATASKLPPPVARKSOLPVYKLLPS 1620  
 1621 QNRLQPKQHVSTFGDDMPRVYCVETGPIINFSTATISLSDLTIESPPNELAAGSGVRGGAQ 1680  
 1621 QNRLQPKQHVSTFGDDMPRVYCVETGPIINFSTATISLSDLTIESPPNELAAGSGVRGGAQ 1680  
 1681 SGFEKEDTITPTGRSTDEAQGGKTSSVTIPELDDNKABEGDILAEICINSAMPKGSHPK 1740  
 1681 SGFEKEDTITPTGRSTDEAQGGKTSSVTIPELDDNKABEGDILAEICINSAMPKGSHPK 1740  
 1741 FRVKIMQVQOASASSAPNKNQLOGKKKPTSPVKPIPONTYRTRVRKNADSKNLN 1800  
 1741 FRVKIMQVQOASASSAPNKNQLOGKKKPTSPVKPIPONTYRTRVRKNADSKNLN 1800  
 1801 AERFSDNKSQKONLKNNSKDFNDKLPNNEDVRGSAFDPSPHYTPIEGTGYCFSRND 1860  
 1801 AERFSDNKSQKONLKNNSKDFNDKLPNNEDVRGSAFDPSPHYTPIEGTGYCFSRND 1860  
 1861 SLSSLDDEDDVDLSREKAEIRKAKENKSEAKVTGHELTSTNOQANKTQAAKQPINR 1920  
 1861 SLSSLDDEDDVDLSREKAEIRKAKENKSEAKVTGHELTSTNOQANKTQAAKQPINR 1920  
 1921 GQPKPILQKOSTPQSSKDI:PDGGAATDEKLFQFAIENTPVCFSHNSLSLSDIDQENN 1980  
 1921 GQPKPILQKOSTPQSSKDI:PDGGAATDEKLFQFAIENTPVCFSHNSLSLSDIDQENN 1980  
 1981 NKENEP:KETEPDPSQGEPSKPOASGYAPKSFHVEDTPVCFSRNSLSLSIDSEDDLQ 2040  
 1981 NKENEP:KETEPDPSQGEPSKPOASGYAPKSFHVEDTPVCFSRNSLSLSIDSEDDLQ 2040  
 2041 ECISAMPKPKKPSRLKGDNEKSPRNMGGILGEDTLDLKDIQRPDSEHGLSPDSENF 2100  
 2041 ECISAMPKPKKPSRLKGDNEKSPRNMGGILGEDTLDLKDIQRPDSEHGLSPDSENF 2100  
 2101 WKATQEGANSIVSLHQAACLSRQASDSDSIISLKSIGLSGSPFHUTPQOEKPF 2160  
 2101 WKATQEGANSIVSLHQAACLSRQASDSDSIISLKSIGLSGSPFHUTPQOEKPF 2160  
 2161 SNKGPRILKPEKXSTLETKKIESKGIKGGKVKYSLITGKVRNSEISQOMKQLOAN 2220  
 2161 SNKGPRILKPEKXSTLETKKIESKGIKGGKVKYSLITGKVRNSEISQOMKQLOAN 2220  
 2221 MPSISRGRTMIHIPGVNNSSTSPVSKGPPKLTTPASKSPSEGTATTSPRGAKPSVK 2280  
 2221 MPSISRGRTMIHIPGVNNSSTSPVSKGPPKLTTPASKSPSEGTATTSPRGAKPSVK 2280  
 2281 ELSVPARQTSIQGSSKAPSRSGRSTPSRPAQOPLSRPIQSPGRNSISPGRNGISPPN 2340  
 2281 ELSVPARQTSIQGSSKAPSRSGRSTPSRPAQOPLSRPIQSPGRNSISPGRNGISPPN 2340  
 2341 KLSQLPRTSSPSTASTKSSGKMSVTSPOGRQMSQNLTKQTGLSKNASSI:PRSESASKG 2400



Db 2341 KLSQPTSTSPSTASTKSSGSKMSYTSPPQWQQQLTKQTGLSKNASSIPSEASKG 2400  
 Qy 2401 LNMNNGNGANKKVELSRMSTKSSGSEDRSPVLVRQSTTIKEAPSTLARKLEESA 2460  
 Db 2401 LNMNNGNGANKKVELSRMSTKSSGSEDRSPVLVRQSTTIKEAPSTLARKLEESA 2460  
 Qy 2461 SFESLSPSPSPSTRSQATPVLSPSLPDMYSLSTHSSVQAGGWRKLPNLSITIIYNDG 2520  
 Db 2461 SFESLSPSPSPSTRSQATPVLSPSLPDMYSLSTHSSVQAGGWRKLPNLSITIIYNDG 2520  
 Qy 2521 RPAKRHDIAHSHSPSLPINSRSGTWKREHSHSSSLPRVSTWRTTGGSSSILSASSES 2580  
 Db 2521 RPAKRHDIAHSHSPSLPINSRSGTWKREHSHSSSLPRVSTWRTTGGSSSILSASSES 2580  
 Qy 2581 SEKAKSEDEKHVNSISGTQSKENQVSAKGTWRKIKENEFSPNTSQTIVSSGATNGAES 2640  
 Db 2581 SEKAKSEDEKHVNSISGTQSKENQVSAKGTWRKIKENEFSPNTSQTIVSSGATNGAES 2640  
 Qy 2641 KTLIYQMAPAVSKTEDVWVRIEDCPINPRSGRSPGTNTPPVDSVSEKANPNIKDSKDN 2700  
 Db 2641 KTLIYQMAPAVSKTEDVWVRIEDCPINPRSGRSPGTNTPPVDSVSEKANPNIKDSKDN 2700  
 Qy 2701 QAKQNVGSGVPMRTVGLNRLNLSFIQVDAPDQKTEIKPGQNNPVPVSETNESSIVERT 2760  
 Db 2701 QAKQNVGSGVPMRTVGLNRLNLSFIQVDAPDQKTEIKPGQNNPVPVSETNESSIVERT 2760  
 Qy 2761 PFSSSSSKHSSPSGTVAARVTPFNYPNPSPRKSADSTARSPOIPTPVNNNTKKEDSKT 2820  
 Db 2761 PFSSSSSKHSSPSGTVAARVTPFNYPNPSPRKSADSTARSPOIPTPVNNNTKKEDSKT 2820  
 Qy 2821 DSTESSGTQSPKRGSGSYLTVS 2843  
 Db 2821 DSTESSGTQSPKRGSGSYLTVS 2843

RESULT 14  
 AAW38370  
 ID AAW38370 standard; protein; 2843 AA.  
 XX AC AAW38370;  
 XX DT 25-MAR-2003 (revised)  
 XX DT 08-APR-1998 (first entry)  
 XX DE Human adenomatous Polyposis coli gene product.  
 XX KW Human; adenomatous Polyposis coli; APC; diagnosis; prognosis;  
 KW neoplastic tissue; tumour tissue; tumour repressor; mutation;  
 KW sporadic colorectal cancer; detection.  
 XX OS Homo sapiens.  
 XX PN US5691454-A.  
 XX PD 25-NOV-1997.  
 XX PF 25-MAY-1995; 95US-00452654.  
 XX PR 16-JAN-1991; 91GB-00000962.  
 XX PR 16-JAN-1991; 91GB-00000963.  
 XX PR 16-JAN-1991; 91GB-00000974.  
 XX PR 16-JAN-1991; 91GB-00000975.  
 XX PR 08-AUG-1991; 91US-00741940.  
 XX PR 12-AUG-1994; 94US-00289548.  
 XX (CANC-) CANCER INST.  
 XX PA (UJO) UNIV JOHNS HOPKINS.  
 XX PA (ICIL) IMPERIAL CHEM IND PLC.  
 XX PA (UTAH) UNIV UTAH.  
 XX Nakamura Y, Markham AP, Groden J, Vogelstein B, Kinzler K,  
 PI Hedge PJ, Carlson M, Anand R, Thliveris A, Albertsen H, White RL;

PI Joslyn G;  
 DR WPI; 1998-017712/02.  
 DR N-PSDB; AAT96153.  
 XX  
 PT Antibodies to normal and mutant adenomatous polyposis coli proteins -  
 useful for detecting genetic predisposition to cancer.  
 XX  
 XX Example 1; Col 33-52; 107pp; English.  
 XX  
 CC The present sequence is the human adenomatous Polyposis coli (APC) gene  
 product, which was used in the development of a novel method of  
 diagnosing or prognosing an APC gene associated neoplastic tissue. The  
 method comprises comparing APC gene coding sequences or mRNA in a tumour  
 tissue, to APC gene coding sequences or mRNA in a non-neoplastic tissue,  
 where a difference indicates an APC gene associated neoplasia of the  
 tumour tissue. APC is a tumour repressor expressed in most normal  
 tissues. APC mutations are found in familial adenomatous polyposis and  
 sporadic colorectal cancer patients. The method enables mutations to be  
 detected to provide an indication of predisposition to cancer. (Updated  
 on 25-MAR-2003 to correct PR field.)  
 XX  
 SQ Sequence 2843 AA;  
 Query Match 99.8%; Score 14533; DB 2; Length 2843;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 2836; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
 Qy 1 MAASDYDQLLKQVAKLKMENSLRQELSDNSNHLTKLETSANNKEVLKQLQGSIEDEAM 60  
 Db 1 MAASDYDQLLKQVAKLKMENSLRQELSDNSNHLTKLETSANNKEVLKQLQGSIEDEAM 60  
 Qy 61 ASSGQIDLLERLKEINLDSNFPFGVKLRKMSLSYSGREGSVSRGEGSPVPMGSPFR 120  
 Db 61 ASSGQIDLLERLKEINLDSNFPFGVKLRKMSLSYSGREGSVSRGEGSPVPMGSPFR 120  
 Qy 121 RGFVNGSRSTGYLELEKERSLLADLDKEEKQWYQALQNLTKRIDSLPTENFSL 180  
 Db 121 RGFVNGSRSTGYLELEKERSLLADLDKEEKQWYQALQNLTKRIDSLPTENFSL 180  
 Qy 181 QTDTRRQLEYEARQIRVAMEEQLGTQDMKRAORRIARIQIEKDIRILOLQSQAT 240  
 Db 181 QTDTRRQLEYEARQIRVAMEEQLGTQDMKRAORRIARIQIEKDIRILOLQSQAT 240  
 Qy 241 EARSQQNKHETGSHDAERQNEGQVGEINMATSNGQGGSTTRMDHETASVLSSTHSA 300  
 Db 241 EARSQQNKHETGSHDAERQNEGQVGEINMATSNGQGGSTTRMDHETASVLSSTHSA 300  
 Qy 301 PRRLTSHLGTVMYVYLLSMGLTHDKDMSRTLLAMSSQDSCISMKGSCPLLIQLL 360  
 Db 301 PRRLTSHLGTVMYVYLLSMGLTHDKDMSRTLLAMSSQDSCISMKGSCPLLIQLL 360  
 Qy 361 HGNDKDSVLLGNSRGSKEARASAAALHNIHSQPDGRGRREIRVHLLEQIRAYCETC 420  
 Db 361 HGNDKDSVLLGNSRGSKEARASAAALHNIHSQPDGRGRREIRVHLLEQIRAYCETC 420  
 Qy 421 WEQBAHFGMDQDNMPAPVEHQICPACVLMKLSFDEHRRHANNELGLOAIELLQ 480  
 Db 421 WEQBAHFGMDQDNMPAPVEHQICPACVLMKLSFDEHRRHANNELGLOAIELLQ 480  
 Qy 481 VDCMYGLTNDHYISITLRRYAGMALTNLTFGDVANKATLCMKGCMALVAQKSEEDL 540  
 Db 481 VDCMYGLTNDHYISITLRRYAGMALTNLTFGDVANKATLCMKGCMALVAQKSEEDL 540  
 Qy 541 QQVIASVLRNLSWRADVNSKKTLEVGSVKALMECALEVKESTLKSIVLSALNLSAHT 600  
 Db 541 QQVIASVLRNLSWRADVNSKKTLEVGSVKALMECALEVKESTLKSIVLSALNLSAHT 600  
 Qy 601 ENKADICAVDQALAFVLTITYSQTNLTALIESGGGILRNVSLSIATNEDHQLRENN 660  
 Db 601 ENKADICAVDQALAFVLTITYSQTNLTALIESGGGILRNVSLSIATNEDHQLRENN 660  
 Qy 661 CLQTLQHLKSHSLTIVSNACGLTNLNLNLSARNPKQDEALWDGAVSMUKNLIHSHKHMAM 720

Db 661 CLOTLLOHLKSHSLTIVSNA CGTLNLSARIPKQOEALWDNGAVSMLXNLIHSRHKMIAM 720  
Qy 721 GSAALRLNLMANRPKAKYKDNIMSGSLPSLHVKQKALBAELDAQHLSLTFNIDNLS 780  
Db 721 GSAALRLNLMANRPKAKYKDNIMSGSLPSLHVKQKALBAELDAQHLSLTFNIDNLS 780  
Qy 781 PKASHRSKORHKQSLGYDVPDTHPHDDNRDNENTGNNTVLSPVLVNTTVLPSSSSRG 840  
Db 781 PKASHRSKORHKQSLGYDVPDTHPHDDNRDNENTGNNTVLSPVLVNTTVLPSSSSRG 840  
Qy 841 LDSSRSKDRSLERERGI GLGNYPHYPATENPGTSSKRGLOI STTAAQIAKVMEEVSALHTS 900  
Db 841 LDSSRSKDRSLERERGI GLGNYPHYPATENPGTSSKRGLOI STTAAQIAKVMEEVSALHTS 900  
Qy 901 QEDRSSGTTLHCVTDERNALRRSSAAHTSNNTNFTKSENSNRTCSMPYAKLEYKRSS 960  
Db 901 QEDRSSGTTLHCVTDERNALRRSSAAHTSNNTNFTKSENSNRTCSMPYAKLEYKRSS 960  
Qy 961 NDLSNSVSSSDGYKRGQMKPSIESYEDDDESKFCYGYQYPADLAHKIHSANHMDNDGE 1020  
Db 961 NDLSNSVSSSDGYKRGQMKPSIESYEDDDESKFCYGYQYPADLAHKIHSANHMDNDGE 1020  
Qy 1021 LDTPNYSIKYSDGOLNSGRQSPSONERWARPKHIIIEIKQSEORQORNOSTTYPVYTE 1080  
Db 1021 LDTPNYSIKYSDGOLNSGRQSPSONERWARPKHIIIEIKQSEORQORNOSTTYPVYTE 1080  
Qy 1081 STDDKHLKFPQHPFGQECVSPVRSRGANGSETNRVGSNHGINQNVQSLSLCOEDDYEDDKP 1140  
Db 1081 STDDKHLKFPQHPFGQECVSPVRSRGANGSETNRVGSNHGINQNVQSLSLCOEDDYEDDKP 1140  
Qy 1141 TNSERYSEEBEHEERPTNYSIKYNEBKRHVDQPIDYSLKYATDIPSSOKQSPFSKS 1200  
Db 1141 TNSERYSEEBEHEERPTNYSIKYNEBKRHVDQPIDYSLKYATDIPSSOKQSPFSKS 1200  
Qy 1201 SGOQSKTHMSSGENTSTPSSNAKRONQLHPSSAQSRGQPOKAATCKYSSINOETIQ 1260  
Db 1201 SGOQSKTHMSSGENTSTPSSNAKRONQLHPSSAQSRGQPOKAATCKYSSINOETIQ 1260  
Qy 1261 TYCVEDTPICFRCSSLSLSSAEDEIGCNQTOQEADANTLIQIAEIKEKIGTRSAEDPV 1320  
Db 1261 TYCVEDTPICFRCSSLSLSSAEDEIGCNQTOQEADANTLIQIAEIKEKIGTRSAEDPV 1320  
Qy 1321 SVPAVQSHPRTKSRRLQSSLSLSSARHKAVERFSSGAKSPSKGAOTPKPPEHYVOET 1380  
Db 1321 SVPAVQSHPRTKSRRLQSSLSLSSARHKAVERFSSGAKSPSKGAOTPKPPEHYVOET 1380  
Qy 1381 PLMFRCSTSVSLDSFESRSIASSVQSPCQMWGSLIIPSDLPDSPGOTMPPSRSKTTP 1440  
Db 1381 PLMFRCSTSVSLDSFESRSIASSVQSPCQMWGSLIIPSDLPDSPGOTMPPSRSKTTP 1440  
Qy 1441 PPPQTAQTKREVPKNKAPTAKRESGPKOAAVNAAVQRVQVLPDADTLLHFATESTPDGF 1500  
Db 1441 PPPQTAQTKREVPKNKAPTAKRESGPKOAAVNAAVQRVQVLPDADTLLHFATESTPDGF 1500  
Qy 1501 SCSSSLSALSLEPFTOKDVELRIMPVQENDNGNETESEOKESENEOEKAEKTI DSE 1560  
Db 1501 SCSSSLSALSLEPFTOKDVELRIMPVQENDNGNETESEOKESENEOEKAEKTI DSE 1560  
Qy 1561 KOLLDSDDDDDIIEILBECIIISAMPTKSSRKAKKPAQTASKLPPPVARPQSLPVYKLLPS 1620  
Db 1561 KOLLDSDDDDDIIEILBECIIISAMPTKSSRKAKKPAQTASKLPPPVARPQSLPVYKLLPS 1620  
Qy 1621 QNRLOQKHVSFTPGDDMPRVYCVETGPIINFSTATSLDLTIESPNELAAEGVRRGAQ 1680  
Db 1621 QNRLOQKHVSFTPGDDMPRVYCVETGPIINFSTATSLDLTIESPNELAAEGVRRGAQ 1680  
Qy 1681 SGEFEKRDITPTEGRSTDEAQGKTSSTVITPELDDNKAEEGDIILAEBCINSAMPKGSHPK 1740  
Db 1681 SGEFEKRDITPTEGRSTDEAQGKTSSTVITPELDDNKAEEGDIILAEBCINSAMPKGSHPK 1740  
Qy 1741 FRVKIMDQVOQASASSAPNKNQLDGKKKPTSPVKPIQONTEYRTRVRKNADSKNNLN 1800

Db 1741 FRVKIMDQVOQASASSAPNKNQLDGKKKPTSPVKPIQONTEYRTRVRKNADSKNNLN 1800  
Qy 1801 AERVFSDNKKDKONLKNNSKOPFNDKLPNNEDVRGSPAFDSPHHYTPIEGTFCYCSERND 1860  
Db 1801 AERVFSDNKKDKONLKNNSKOPFNDKLPNNEDVRGSPAFDSPHHYTPIEGTFCYCSERND 1860  
Qy 1861 SLSSLPDDDDVDLSREKAEELRAKENKSEAKVTSHTELTSNQOQSAKNTQAIKAPINR 1920  
Db 1861 SLSSLPDDDDVDLSREKAEELRAKENKSEAKVTSHTELTSNQOQSAKNTQAIKAPINR 1920  
Qy 1921 GQKXPILQKOSTPPOSQSKOIPDRGAATDEKLONFJAIENTPVCFSHNSLSLSDIOENN 1980  
Db 1921 GQKXPILQKOSTPPOSQSKOIPDRGAATDEKLONFJAIENTPVCFSHNSLSLSDIOENN 1980  
Qy 1981 NKNEPIKETEPEDSOQEPSPKQASGYAPKSPHVEDTPVCFSRNSLSLSIDSEDDLQ 2040  
Db 1981 NKNEPIKETEPEDSOQEPSPKQASGYAPKSPHVEDTPVCFSRNSLSLSIDSEDDLQ 2040  
Qy 2041 ECISAPVKKKPSRLKGDNEKHSPRNMGGILGEDITLLOKJQRPDSEHGLSPDSENF 2100  
Db 2041 ECISAPVKKKPSRLKGDNEKHSPRNMGGILGEDITLLOKJQRPDSEHGLSPDSENF 2100  
Qy 2101 WKAIQEGANSIVSLHQAAAAACLSRQAASDSDSILSLKSGISLGSPPHLPDQEKFP 2160  
Db 2101 WKAIQEGANSIVSLHQAAAAACLSRQAASDSDSILSLKSGISLGSPPHLPDQEKFP 2160  
Qy 2161 SNKGPRILKFGKSTLTETKKIESESGIKGKKVYKSLITGKYRSNSEISGQMKQPLQAN 2220  
Db 2161 SNKGPRILKFGKSTLTETKKIESESGIKGKKVYKSLITGKYRSNSEISGQMKQPLQAN 2220  
Qy 2221 MPISIRGRITWIIHPGVNNSSTSPVSKGPPILKTPASKSPSEGOTATTSRGAKS VKS 2280  
Db 2221 MPISIRGRITWIIHPGVNNSSTSPVSKGPPILKTPASKSPSEGOTATTSRGAKS VKS 2280  
Qy 2281 ELSFVARQTSQIOGGSSKAPSRSGSRDSTPSRPAQQLSRPIQSPGRNSIPGRNGISPPN 2340  
Db 2281 ELSFVARQTSQIOGGSSKAPSRSGSRDSTPSRPAQQLSRPIQSPGRNSIPGRNGISPPN 2340  
Qy 2341 KLSOLPTSPSTASTKSSGSGKMSYTSPOQVSOQNLTKQTLGSKXASSIPRESASKG 2400  
Db 2341 KLSOLPTSPSTASTKSSGSGKMSYTSPOQVSOQNLTKQTLGSKXASSIPRESASKG 2400  
Qy 2401 LNMNNGNGANKKVELSRMSSTKSSGSDSRPVRPVLVROSTFIKAPSTTLRKLKLESA 2460  
Db 2401 LNMNNGNGANKKVELSRMSSTKSSGSDSRPVRPVLVROSTFIKAPSTTLRKLKLESA 2460  
Qy 2461 SFESLSPSSRPASPTRSQAQTPVLSPLDMSLSTHSSVOAGWRKLPPLNLSPTIEYNDG 2520  
Db 2461 SFESLSPSSRPASPTRSQAQTPVLSPLDMSLSTHSSVOAGWRKLPPLNLSPTIEYNDG 2520  
Qy 2521 RPAKRHDIAKSHSPPSRLPINRSGTWKREHSHKSSSLPRVSTWRRPTGSSSILSASSES 2580  
Db 2521 RPAKRHDIAKSHSPPSRLPINRSGTWKREHSHKSSSLPRVSTWRRPTGSSSILSASSES 2580  
Qy 2581 SEKAKSEDEKHVNSISGTSKSKENQVSAKGTWRKIKENEFSPNTNSTQTVSSGATNGAES 2640  
Db 2581 SEKAKSEDEKHVNSISGTSKSKENQVSAKGTWRKIKENEFSPNTNSTQTVSSGATNGAES 2640  
Qy 2641 KTLIYQMAPAVSKTDEWVRIEDCPINNPRSGRSPGTQTPPVIDSVEKAPNPKOSKDN 2700  
Db 2641 KTLIYQMAPAVSKTDEWVRIEDCPINNPRSGRSPGTQTPPVIDSVEKAPNPKOSKDN 2700  
Qy 2701 QAKQNVGSGVPMTVGLNRLNSFIQVDAPDOKGTETIKPGQNNPVVSVSTNESSI VERT 2760  
Db 2701 QAKQNVGSGVPMTVGLNRLNSFIQVDAPDOKGTETIKPGQNNPVVSVSTNESSI VERT 2760  
Qy 2761 PFSSSSSKHSSPSGTTVAARVTPPNYNSPRKSSADTSARPQIPTPVNNNTKKRUSKT 2820  
Db 2761 PFSSSSSKHSSPSGTTVAARVTPPNYNSPRKSSADTSARPQIPTPVNNNTKKRUSKT 2820  
Qy 2821 DSTESSGTSQSPKRHSGSYLVTSV 2843  
Db 2821 DSTESSGTSQSPKRHSGSYLVTSV 2843

RESULT 15

ABG90964

ID ABG90964 standard; protein; 2843 AA.

XX

AC ABG90964;

XX 29-NOV-2002 (first entry)

XX Human adenomatous polyposis coli (APC) protein #1.

XX Adenomatous polyposis coli; APC; human; neoplastic tissue;

XX mutation detection; tumour; cancer.

XX Homo sapiens.

XX OS

XX US6413727-B1.

XX PD

XX 02-JUL-2002.

XX PF

XX 25-MAY-1995; 95US-00449731.

XX PR

XX 16-JAN-1991; 91GB-00000962.

XX PR

XX 16-JAN-1991; 91GB-00000963.

XX PR

XX 16-JAN-1991; 91GB-00000974.

XX PR

XX 16-JAN-1991; 91GB-00000975.

XX PR

XX 08-AUG-1991; 91US-00741940.

XX PR

XX 12-AUG-1994; 94US-00289548.

XX XX

XX (UXJO ) UNIV JOHNS HOPKINS.

XX PA

XX (UTAH ) UNIV UTAH.

XX PA

XX (NICA-) JAPANESE FOUND CANCER RES.

XX PA

XX (ZENE ) ZENECA LTD.

XX XX

XX Albertsen H, Anand R, Carlson M, Groden J, Hedge PU, Joslyn G;

XX PI

XX Kinzler K, Markham AF, Nakamura Y, Thliveris A, Vogelstein B;

XX PI

XX White RL;

XX XX

XX WPI: 2002-641559/69.

XX DR

XX N-PSDB; ABS67119.

XX DR

XX Method to aid in the diagnosis/prognosis of neoplastic tissues in humans,

XX PT

XX by detecting somatic alteration of wild-type APC protein in tumor tissue

XX PT

XX isolated from human, the alteration indicating neoplasia of the tissue.

XX XX

XX Disclosure; Fig 3; 140pp; English.

XX PS

XX This invention relates to a novel method to aid in the diagnosis or

XX CC

XX prognosis of a neoplastic tissue of a human. The method involves

XX CC

XX detecting somatic alteration of wild-type adenomatous polyposis coli)

XX CC

XX protein in a tumour tissue isolated from a human (the alteration

XX CC

XX indicating neoplasia of the tissue). The method of the invention is

XX CC

XX useful in diagnosis or prognosis of a neoplastic tissue of a human. the

XX CC

XX method is useful in detection of genetic predisposition to cancer. The

XX CC

XX present sequence represents a protein sequence used in the method of the

XX CC

XX invention

XX XX

XX Sequence 2843 AA;

XX SQ

XX Query Match

XX Best Local Similarity 99.8%; Score 14533; DB 5; Length 2843;

XX Matches 2836; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

XX

XX 1 MAAASVDQLLKQVEALKMENSNLRLQEDNSNHLTKLETEASNKKEVLKQLQGSIEDEAM 60

XX DB

XX 1 MAAASVDQLLKQVEALKMENSNLRLQEDNSNHLTKLETEASNKKEVLKQLQGSIEDEAM 60

XX QY

XX 61 ASSGQIDLLRLKELNLDSSNFFGVKLRKMSLRYSRGSGVSRSGCSVPVPGSPFR 120

XX DB

XX 61 ASSGQIDLLRLKELNLDSSNFFGVKLRKMSLRYSRGSGVSRSGCSVPVPGSPFR 120

XX QY

XX 121 RGFVNGSRSTGYLEELEKERSLLADLDKEEKEKWYQAQLNLTKRIDSPLTENFSL 180

XX

XX 121 RGFVNGSRSTGYLEELEKERSLLADLDKEEKEKWYQAQLNLTKRIDSPLTENFSL 180

XX

XX 181 CTDMTRRQLEVEARQIRVAMEEQIGTODMEKRAQRRIARIQQIEKDIRLOLQSOAT 240

XX QY

XX 181 CTDMTRRQLEVEARQIRVAMEEQIGTODMEKRAQRRIARIQQIEKDIRLOLQSOAT 240

XX DB

XX 181 CTDMTRRQLEVEARQIRVAMEEQIGTODMEKRAQRRIARIQQIEKDIRLOLQSOAT 240

XX QY

XX 241 EAERSSQNKHETGSHDAERQNEGQGVGEINMATSGNGQSTTRMDHETASVLSSTHSA 300

XX DB

XX 241 EAERSSQNKHETGSHDAERQNEGQGVGEINMATSGNGQSTTRMDHETASVLSSTHSA 300

XX QY

XX 301 PRRLTSHLGTKEVYVYSLLSMLGTHDKDDNSRTLLAMSSQSDSCISMRQSCPLLIQLL 360

XX DB

XX 301 PRRLTSHLGTKEVYVYSLLSMLGTHDKDDNSRTLLAMSSQSDSCISMRQSCPLLIQLL 360

XX QY

XX 361 HGNDKDSVLLGNRSRGSKEARARASAAALHNIHSOPDDKRGREIRVLHLLLEQIRAYCETC 420

XX DB

XX 361 HGNDKDSVLLGNRSRGSKEARARASAAALHNIHSOPDDKRGREIRVLHLLLEQIRAYCETC 420

XX QY

XX 421 WEMQEAHEPGMDQDKNMPAPVEHQICPAVCVLMKLSFDEEHRHAMNELGGLQIAELLQ 480

XX DB

XX 421 WEMQEAHEPGMDQDKNMPAPVEHQICPAVCVLMKLSFDEEHRHAMNELGGLQIAELLQ 480

XX QY

XX 481 VDCMYGLTNDHYSTITLRRYAGVATLNLTFGDVANKATLCMKGCMRALVAQLKSEEDL 540

XX DB

XX 481 VDCMYGLTNDHYSTITLRRYAGVATLNLTFGDVANKATLCMKGCMRALVAQLKSEEDL 540

XX QY

XX 541 QQVIASVLRNLSWRADVNSKKTLEVGSVKALMECALEVKKESTLKSVALNLSAHCT 600

XX DB

XX 541 QQVIASVLRNLSWRADVNSKKTLEVGSVKALMECALEVKKESTLKSVALNLSAHCT 600

XX QY

XX 601 ENKADICAVDGCALAFVLTLYRSQNTNTLAIIESGGGILRNVSLLIATNDDHRLRENN 660

XX DB

XX 601 ENKADICAVDGCALAFVLTLYRSQNTNTLAIIESGGGILRNVSLLIATNDDHRLRENN 660

XX QY

XX 661 CLQTLQHLKSHSLTIVSNACGTLNLSARNPKDQEAALWDMGAVSMKLNLIHSHKHMIAM 720

XX DB

XX 661 CLQTLQHLKSHSLTIVSNACGTLNLSARNPKDQEAALWDMGAVSMKLNLIHSHKHMIAM 720

XX QY

XX 721 GSAAALRNLMANRPAPKYDANIMSPGSSLSLHVYKOKALEAEALDAHLSETFNIDNLS 780

XX DB

XX 721 GSAAALRNLMANRPAPKYDANIMSPGSSLSLHVYKOKALEAEALDAHLSETFNIDNLS 780

XX QY

XX 781 PKASHRSQRHQRKQSLYGDYVFDNRRHDDNSDNFTGNMTVLSPLYNTTVLPSSSSSRGS 840

XX DB

XX 781 PKASHRSQRHQRKQSLYGDYVFDNRRHDDNSDNFTGNMTVLSPLYNTTVLPSSSSSRGS 840

XX QY

XX 841 LDSRSERKDSRLERERIGLGNVHPATENPGTSSKRGLOISTTAAQIAKYMEEVSAIHTS 900

XX DB

XX 841 LDSRSERKDSRLERERIGLGNVHPATENPGTSSKRGLOISTTAAQIAKYMEEVSAIHTS 900

XX QY

XX 901 QEDRSSSTTELCHVTDERVALRPSAAHNSHTNFTKSENSNRTCSMPYAKLEYKRSS 960

XX DB

XX 901 QEDRSSSTTELCHVTDERVALRPSAAHNSHTNFTKSENSNRTCSMPYAKLEYKRSS 960

XX QY

XX 961 NDSLNSVSSSDGYGKRGQMKPSIESYSEDDSKPCSYGYPADIAHKSANHMDNDGE 1020

XX DB

XX 961 NDSLNSVSSSDGYGKRGQMKPSIESYSEDDSKPCSYGYPADIAHKSANHMDNDGE 1020

XX QY

XX 1021 LDTPINYSKYSDQALNSGRSQSPQNERWARPKHIIIEDEIKQSEQRQSGTTPYVTE 1080

XX DB

XX 1021 LDTPINYSKYSDQALNSGRSQSPQNERWARPKHIIIEDEIKQSEQRQSGTTPYVTE 1080

XX QY

XX 1081 STDDKHLKFPFHQQQECVSPYRSRGANGSETNRVGNHGINQNVSLQCEDDYEDDKP 1140

XX DB

XX 1081 STDDKHLKFPFHQQQECVSPYRSRGANGSETNRVGNHGINQNVSLQCEDDYEDDKP 1140

XX QY

XX 1141 TNYSERVSEEEQHEEERPTNYSIKYNEEKRVHVDOPIDYSIKYATDIPSSOKSEFSKS 1200

XX DB

XX 1141 TNYSERVSEEEQHEEERPTNYSIKYNEEKRVHVDOPIDYSIKYATDIPSSOKSEFSKS 1200

XX QY

XX 1201 SSGOSSKTEHMSSESTTSPSSNAKQNLHPSSAQSRGQPKAATCKVSSINQFTIQ 1260

Db 1201 SSGQSKTEHSSSENTSTFSSNAKRONQLHPSSAQSRSGQPQKAATCKVSSINQETIQ 1260  
QY 1261 TYCVEDTPICFSRCSLSLSAEDIEICNTOBADSANTLOIAEIKEKIGITSAEDPV 1320  
Db 1261 TYCVEDTPICFSRCSLSLSAEDIEICNTOBADSANTLOIAEIKEKIGITSAEDPV 1320  
QY 1321 SEVPASQHPRTKSRRLQCSLSSESARHKAVERFSSGAKSPKSGAOPPKSPPHYVOET 1380  
Db 1321 SEVPASQHPRTKSRRLQCSLSSESARHKAVERFSSGAKSPKSGAOPPKSPPHYVOET 1380  
QY 1381 PLMFRCSTSVSLDSFESRSTASSVQSPSCGWSGIIISPSDLDPSPGQOTWPPRSKTPP 1440  
Db 1381 PLMFRCSTSVSLDSFESRSTASSVQSPSCGWSGIIISPSDLDPSPGQOTWPPRSKTPP 1440  
QY 1441 PPPQTAQTKREVPKNKAPTAEKRESGPKQAAVNAAVQVQVLPDADTLHLHPATESTDGF 1500  
Db 1441 PPPQTAQTKREVPKNKAPTAEKRESGPKQAAVNAAVQVQVLPDADTLHLHPATESTDGF 1500  
QY 1501 SCSSLSLSALDEPFIQXDELIMPPVOENDNGNETESEOPKESNENQEAKEKTIDSE 1560  
Db 1501 SCSSLSLSALDEPFIQXDELIMPPVOENDNGNETESEOPKESNENQEAKEKTIDSE 1560  
QY 1561 KDLDDSDDDDIIEILEECIIISAMPTKSSRKAKPAQTASKLPPPVARKPSQLPVVKLLPS 1620  
Db 1561 KDLDDSDDDDIIEILEECIIISAMPTKSSRKAKPAQTASKLPPPVARKPSQLPVVKLLPS 1620  
QY 1621 QNRLOPQKHVSFTPDGDDMPRVYCVEGTPIINFSTATSLDLTIESPPNELAAGEVGRGAQ 1680  
Db 1621 QNRLOPQKHVSFTPDGDDMPRVYCVEGTPIINFSTATSLDLTIESPPNELAAGEVGRGAQ 1680  
QY 1681 SEFEKRDITIPEGRSTDEAOCCKTSSVTIPELDDNKABEGDIIAECINSAMPKGSHP 1740  
Db 1681 SEFEKRDITIPEGRSTDEAOCCKTSSVTIPELDDNKABEGDIIAECINSAMPKGSHP 1740  
QY 1741 FRVKIMDQVQOASASSAPNKNQLDGKKKPTSPVKPIPQNTYRTRVRKNVADSKNLN 1800  
Db 1741 FRVKIMDQVQOASASSAPNKNQLDGKKKPTSPVKPIPQNTYRTRVRKNVADSKNLN 1800  
QY 1801 AERFSDNKKONLKNNSKDFNDKLPNNEDRVGSAFDPSPHHYTHPIEGTPIYCFSRND 1860  
Db 1801 AERFSDNKKONLKNNSKDFNDKLPNNEDRVGSAFDPSPHHYTHPIEGTPIYCFSRND 1860  
QY 1861 SLSSLDFFDDDDVLDLREKAELKAKENKESAKVTSHTELTSNQOSANKTQAIKQPINR 1920  
Db 1861 SLSSLDFFDDDDVLDLREKAELKAKENKESAKVTSHTELTSNQOSANKTQAIKQPINR 1920  
QY 1921 GQPKPILQKQSTFPOSSKDIIPRGAATDEKLQNFALIENTPVCFSNSSLSDIDQENN 1980  
Db 1921 GQPKPILQKQSTFPOSSKDIIPRGAATDEKLQNFALIENTPVCFSNSSLSDIDQENN 1980  
QY 1981 NKENPPIKETEPDQGBPSKPOASGYAPKSPHVEDTPVCFSRNSSLSLSIDSDDLQ 2040  
Db 1981 NKENPPIKETEPDQGBPSKPOASGYAPKSPHVEDTPVCFSRNSSLSLSIDSDDLQ 2040  
QY 2041 ECISAMPKKKPSRLKGDNEKHSPRNMGIIIGEDTLTDLKDIOQPDSEHGLSPDSENF 2100  
Db 2041 ECISAMPKKKPSRLKGDNEKHSPRNMGIIIGEDTLTDLKDIOQPDSEHGLSPDSENF 2100  
QY 2101 WKAIQEGANSIVSSLHOAAAACLRQASSDSDLSLSKSGISLSPHLPDQEEKPT 2160  
Db 2101 WKAIQEGANSIVSSLHOAAAACLRQASSDSDLSLSKSGISLSPHLPDQEEKPT 2160  
QY 2161 SNKGPRILKPGKSTLTETKIESESGIKGGKVKYKSLITGVRNSNISEGQMKQPLQAN 2220  
Db 2161 SNKGPRILKPGKSTLTETKIESESGIKGGKVKYKSLITGVRNSNISEGQMKQPLQAN 2220  
QY 2221 MFSISGRWTMIHPGVRNSSSTSPVSKGPKLPKTPASKSPSEGQTATTSRGAKPSVKS 2280  
Db 2221 MFSISGRWTMIHPGVRNSSSTSPVSKGPKLPKTPASKSPSEGQTATTSRGAKPSVKS 2280  
QY 2281 ELSPVARQTSQIGGSSKAPSRGSRDSTPSRPAQOPLSRPIOSPGRNSTSPGRNGISPPN 2340  
Db 2281 ELSPVARQTSQIGGSSKAPSRGSRDSTPSRPAQOPLSRPIOSPGRNSTSPGRNGISPPN 2340

QY 2341 KLSQLPRTSPSTASTKSSGSGKMSYTSFGRQMSQONLTQOTGLSKNASSIPRSESASKG 2400  
Db 2341 KLSQLPRTSPSTASTKSSGSGKMSYTSFGRQMSQONLTQOTGLSKNASSIPRSESASKG 2400  
QY 2401 LNOVNGNGANKKVELSRMSSSTKSSGSESDRSPVLRQSTFIKEAPSPTLERKLEESA 2460  
Db 2401 LNOVNGNGANKKVELSRMSSSTKSSGSESDRSPVLRQSTFIKEAPSPTLERKLEESA 2460  
QY 2461 SFESLSPSGRPASPTRSQAQTPVLSPLPDMSLSTHSSVQAGGWRKLPNLSPTIEYNDG 2520  
Db 2461 SFESLSPSGRPASPTRSQAQTPVLSPLPDMSLSTHSSVQAGGWRKLPNLSPTIEYNDG 2520  
QY 2521 RPAKRHDIAKSHSESPSRPIPNRSGTWKREHSHSSSLPRVSTWRTTGSSSSILSASSES 2580  
Db 2521 RPAKRHDIAKSHSESPSRPIPNRSGTWKREHSHSSSLPRVSTWRTTGSSSSILSASSES 2580  
QY 2581 SEKAKSEDEKHVNSISGTKQSKENOVSAKGTWRKIKENEFSPNTNSTSQTVSSGATNGAES 2640  
Db 2581 SEKAKSEDEKHVNSISGTKQSKENOVSAKGTWRKIKENEFSPNTNSTSQTVSSGATNGAES 2640  
QY 2641 KTLIYQMAPAVSKTEDVWVRIEDCPINNPRSGRSPGTGNTPPVJDSVSEKANPNIKOSKDN 2700  
Db 2641 KTLIYQMAPAVSKTEDVWVRIEDCPINNPRSGRSPGTGNTPPVJDSVSEKANPNIKOSKDN 2700  
QY 2701 QAKONTYGVGSPVWRTVGLNRLNSFIQVADAPDOCKTEIKPGQNNPVPVSETNESSIVERT 2760  
Db 2701 QAKONTYGVGSPVWRTVGLNRLNSFIQVADAPDOCKTEIKPGQNNPVPVSETNESSIVERT 2760  
QY 2761 PFSSSSSSKSHSSPSGTVAARVTPFNYNPSPRKSSADSTARPQOIPTPVNNNTKKRDSKT 2820  
Db 2761 PFSSSSSSKSHSSPSGTVAARVTPFNYNPSPRKSSADSTARPQOIPTPVNNNTKKRDSKT 2820  
QY 2821 DSTESSGTOQPKHSGSYLTVSV 2843  
Db 2821 DSTESSGTOQPKHSGSYLTVSV 2843

Search completed: August 25, 2004, 17:15:10

Job time : 195.5 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2004, 17:23:12 ; Search time 190.5 Seconds  
(without alignments)  
4695.247 Million cell updates/sec

Title: US-09-442-489f-7  
Perfect score: 14566  
Sequence: 1 MAASVDLLKQVEALKNEN.....ESSGTQSPKRGSGYLTVS 2843

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1297172 seqs, 314612898 residues

Total number of hits satisfying chosen parameters: 1297172

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US05\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09D\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10E\_PUBCOMB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US10F\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	14566	100.0	2843	9	US-09-987-482-1
2	14566	100.0	2843	12	US-10-392-113-21
3	14557	99.9	2843	16	US-10-408-765A-1970
4	14539	99.8	2843	8	US-08-681-219-32
5	14539	99.8	2843	11	US-09-230-111C-30
6	14539	99.8	2843	14	US-10-092-138-30
7	14484.5	99.4	2844	12	US-10-267-502-370
8	13178	90.5	2845	12	US-10-267-502-372
9	4738	32.5	912	9	US-09-987-482-2
10	3892	26.7	767	9	US-09-987-482-3
11	3512.5	24.1	2274	12	US-10-267-502-373
12	3415.5	23.4	2303	12	US-10-267-502-371
13	1765	12.1	2417	12	US-10-267-502-368
14	1652	11.3	325	12	US-09-915-307-5
15	1477	10.1	1633	14	US-10-029-386-33090

16	1222.5	8.4	1067	12	US-10-267-502-369	Sequence 369, Appl
17	562.5	3.9	2344	9	US-09-915-242-12713	Sequence 12713, A
18	536	3.7	2263	16	US-10-408-765A-2231	Sequence 2231, Ap
19	517	3.5	2271	12	US-10-282-122A-43924	Sequence 43924, A
20	516	3.5	2283	14	US-10-172-502-4	Sequence 4, Appli
21	510	3.5	4322	16	US-10-437-963-104793	Sequence 104793, A
22	507.5	3.5	178	13	US-10-071-751-49	Sequence 49, Appl
23	474	3.3	5560	12	US-10-263-929-142	Sequence 142, App
24	463.5	3.2	6641	12	US-10-282-122A-70580	Sequence 70580, A
25	462.5	3.2	10203	16	US-10-661-809-23	Sequence 23, Appl
26	454	3.1	3507	15	US-10-369-493-5784	Sequence 5784, Ap
27	453	3.1	4776	12	US-10-282-122A-73678	Sequence 73678, A
28	452	3.1	3664	12	US-10-263-929-143	Sequence 143, App
29	452	3.1	3664	14	US-10-177-293-423	Sequence 423, App
30	452	3.1	3664	16	US-10-408-765A-2287	Sequence 2287, Ap
31	447.5	3.1	22152	16	US-10-715-066-5	Sequence 5, Appli
32	443.5	3.0	1831	12	US-10-282-122A-71033	Sequence 71033, A
33	436	3.0	2655	9	US-09-864-761-34248	Sequence 34248, A
34	430.5	3.0	3551	12	US-10-263-929-144	Sequence 144, App
35	424.5	2.9	2701	14	US-10-171-311-83	Sequence 83, Appl
36	423	2.9	5877	15	US-10-142-515-11	Sequence 11, Appl
37	418	2.9	4723	15	US-10-359-012-8	Sequence 8, Appli
38	415.5	2.9	2803	12	US-10-415-187-5	Sequence 5, Appli
39	413	2.8	3692	12	US-10-282-122A-71235	Sequence 71235, A
40	409.5	2.8	3051	15	US-10-144-194A-62	Sequence 62, Appl
41	408	2.8	5935	14	US-10-243-243A-8	Sequence 8, Appli
42	404.5	2.8	6281	9	US-09-915-242-12996	Sequence 12996, A
43	399	2.7	2137	12	US-10-042-865-81	Sequence 81, Appl
44	399	2.7	2432	12	US-10-112-944-359	Sequence 359, App
45	399	2.7	2545	12	US-10-042-865-12	Sequence 12, Appl

## ALIGNMENTS

RESULT 1  
US-09-987-482-1  
; Sequence 1, Application US/09987482  
; Publication No. US20020184656A1  
; GENERAL INFORMATION:  
; APPLICANT: BHANDARI, POONAM  
; TITLE OF INVENTION: IN VIVO ASSAY SYSTEM FOR SCREENING AND VALIDATION OF  
; TITLE OF INVENTION: DRUGS AND OTHER SUBSTANCES  
; FILE REFERENCE: 056859-0134  
; CURRENT APPLICATION NUMBER: US/09/987,482  
; CURRENT FILING DATE: 2002-03-21  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 2843  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-987-482-1

Query Match	100.0%;	Score 14566;	DB 9;	Length 2843;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2843;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MAASVDLLKQVEALKNENRQLEDNSNLTKLETSNNKVKLQOGSIEDEAM	60	
Db	1	MAASVDLLKQVEALKNENRQLEDNSNLTKLETSNNKVKLQOGSIEDEAM	60	
Qy	61	ASSGQIDLLERLKNELNDSNFFGVKLRSMVLSYSGSREGSVSSRSGECSVPVMSGFPR	120	
Db	61	ASSGQIDLLERLKNELNDSNFFGVKLRSMVLSYSGSREGSVSSRSGECSVPVMSGFPR	120	
Qy	121	RGFVNGSRESTGYLELEKERSLLDADLKEEKEKWKYLAQLNLTNRDLSLTENFSL	180	
Db	121	RGFVNGSRESTGYLELEKERSLLDADLKEEKEKWKYLAQLNLTNRDLSLTENFSL	180	
Qy	181	QTDWTRRLYEALRVAMEEOLGTCCDMKEAKQARIQIOIEKDIRLROLLOSOAT	240	

Db 181 QTDYTRQLYEYARQIRVAMEEQLGTCDQVEKQARQIARIQOIEKDIIRIQLLOSQAT 240  
Qy 241 EAPRSSQNHETGSHDAERQNEQGVGEINMATSNGQGSITRMDHETASVJSSSSTHSA 300  
Db 241 EAPRSSQNHETGSHDAERQNEQGVGEINMATSNGQGSITRMDHETASVJSSSSTHSA 300  
Qy 301 PRRLTSHLGTKEWYVYLLSMLGTHDKDDMSRTLLAASSQDSQSCISMRQSGCLPLLIQLL 360  
Db 301 PRRLTSHLGTKEWYVYLLSMLGTHDKDDMSRTLLAASSQDSQSCISMRQSGCLPLLIQLL 360  
Qy 361 HGNDKDSVLLGNRSGKEARASAAALHNIHSPQDDKGRREIRVHLLEQIRAYCETC 420  
Db 361 HGNDKDSVLLGNRSGKEARASAAALHNIHSPQDDKGRREIRVHLLEQIRAYCETC 420  
Qy 421 WEQEAEHPGMDQKNPMPAPVEHQICPAVCVLMKLSFDEEHRHANEVLGGIOATAELLO 480  
Db 421 WEQEAEHPGMDQKNPMPAPVEHQICPAVCVLMKLSFDEEHRHANEVLGGIOATAELLO 480  
Qy 481 VDCMYGLTNDHYSITLRRVAGMALTNLTFGDVANKATLCSMKGCMRALVAQLKSESDL 540  
Db 481 VDCMYGLTNDHYSITLRRVAGMALTNLTFGDVANKATLCSMKGCMRALVAQLKSESDL 540  
Qy 541 QOVIASVLRNLSWRADVNSKTLREVGSVKALMECALEVKESTLKSVALWNLSAHCT 600  
Db 541 QOVIASVLRNLSWRADVNSKTLREVGSVKALMECALEVKESTLKSVALWNLSAHCT 600  
Qy 601 ENKADICAVGALAFUVTLYRSQNTNLAIIESGGILRNVSLLIATNEDHQILRENN 660  
Db 601 ENKADICAVGALAFUVTLYRSQNTNLAIIESGGILRNVSLLIATNEDHQILRENN 660  
Qy 661 CLQTLQLKSHSLTIVSNACGLTWNLSARNPKDQBALMDMGAVSMLKNLIHSHKHQIAM 720  
Db 661 CLQTLQLKSHSLTIVSNACGLTWNLSARNPKDQBALMDMGAVSMLKNLIHSHKHQIAM 720  
Qy 721 GSAALANLWANRPKADANIMSPGSLPSIAVRKQKALEAELDAQHILSETFDNIDNLS 780  
Db 721 GSAALANLWANRPKADANIMSPGSLPSIAVRKQKALEAELDAQHILSETFDNIDNLS 780  
Qy 781 PKASHRSKQKHQKSLYDGYVDFNRHDDNRSDFNTGNMTVLSPYLTNTVLPSSSSSRGS 840  
Db 781 PKASHRSKQKHQKSLYDGYVDFNRHDDNRSDFNTGNMTVLSPYLTNTVLPSSSSSRGS 840  
Qy 841 LQSSRSKQSLERBERGIGLGNTHPATENPGTSKRGLOITSTAAQIAKWEVSAIHTS 900  
Db 841 LQSSRSKQSLERBERGIGLGNTHPATENPGTSKRGLOITSTAAQIAKWEVSAIHTS 900  
Qy 901 QEDRSSGTTLHCVTDERNALRRSSAAHSHNTYNTFKSENENRCSMPYAKLEYKRSS 960  
Db 901 QEDRSSGTTLHCVTDERNALRRSSAAHSHNTYNTFKSENENRCSMPYAKLEYKRSS 960  
Qy 961 NDSLNVSSSDGYKRGQMKPSEIESYEDDESKFCSYQGYPADLAHKIHSANHMDNDGE 1020  
Db 961 NDSLNVSSSDGYKRGQMKPSEIESYEDDESKFCSYQGYPADLAHKIHSANHMDNDGE 1020  
Qy 1021 LQTPINYSKYSDQLNSGRQSPQNERWARPKHIIEDELKQSEQRNQSTTYPVYTE 1080  
Db 1021 LQTPINYSKYSDQLNSGRQSPQNERWARPKHIIEDELKQSEQRNQSTTYPVYTE 1080  
Qy 1081 STDDKHLKFOFHFQOQECVSPYRSRANGSETNRVGNHGINQVNSQSLCQEDDYEDDKP 1140  
Db 1081 STDDKHLKFOFHFQOQECVSPYRSRANGSETNRVGNHGINQVNSQSLCQEDDYEDDKP 1140  
Qy 1141 TNYSERYSEBQHEEERPTNYSIKYNEEKHVDQPIDYSLKYATDIPSSKQSFSPKS 1200  
Db 1141 TNYSERYSEBQHEEERPTNYSIKYNEEKHVDQPIDYSLKYATDIPSSKQSFSPKS 1200  
Qy 1201 SSGQSKTEHMSSESSENTSTPSSNAKQNLHPSSAQSRSGQPKAATCKVSSINQETIQ 1260  
Db 1201 SSGQSKTEHMSSESSENTSTPSSNAKQNLHPSSAQSRSGQPKAATCKVSSINQETIQ 1260  
Qy 1261 TYCVEDTPICFSRCSLSLSAABDEIGCNQTTQEADSAANTLOIAIEKEKIGTRSAEDPV 1320  
Db 1261 TYCVEDTPICFSRCSLSLSAABDEIGCNQTTQEADSAANTLOIAIEKEKIGTRSAEDPV 1320

Qy 1321 SEVPAVSOHPRTKSRRLQGSLSSESARHKAVERPSSCAKSPSKSGAOTPKSPPEHYVOET 1380  
Db 1321 SEVPAVSOHPRTKSRRLQGSLSSESARHKAVERPSSCAKSPSKSGAOTPKSPPEHYVOET 1380  
Qy 1381 PLMFBSRCTSVSSLDSPFESRSIASSVQSEPCSGMVSGIISPSDLPPSPQTMPPSRSKTTP 1440  
Db 1381 PLMFBSRCTSVSSLDSPFESRSIASSVQSEPCSGMVSGIISPSDLPPSPQTMPPSRSKTTP 1440  
Qy 1441 PPQTAQTKRVVPKNAKAPTAKRESGPKQAANAARVQVLPDADTLLHFAFESTPDGF 1500  
Db 1441 PPQTAQTKRVVPKNAKAPTAKRESGPKQAANAARVQVLPDADTLLHFAFESTPDGF 1500  
Qy 1501 SCSSLSALSDEFFIOKDVELRIMPPVQENDNGNETESQPKESNENQEKAEKTIJSE 1560  
Db 1501 SCSSLSALSDEFFIOKDVELRIMPPVQENDNGNETESQPKESNENQEKAEKTIJSE 1560  
Qy 1561 KDLDDSDDDDIIELEBIIISAMPTKSSRAKAPQATASKLPPPVARKPSQLPVYKLLPS 1620  
Db 1561 KDLDDSDDDDIIELEBIIISAMPTKSSRAKAPQATASKLPPPVARKPSQLPVYKLLPS 1620  
Qy 1621 QNRLOPQKHVSFTPGDDMPRVYCVEGTPIINFSTATSLDLTIESPPNELAAGEGVRGAQ 1680  
Db 1621 QNRLOPQKHVSFTPGDDMPRVYCVEGTPIINFSTATSLDLTIESPPNELAAGEGVRGAQ 1680  
Qy 1681 SGEFEKRDITPTEGRSTDEAQQGKTSSVTIPELDDNKAEBGDILAEICINSAMPKGSHPK 1740  
Db 1681 SGEFEKRDITPTEGRSTDEAQQGKTSSVTIPELDDNKAEBGDILAEICINSAMPKGSHPK 1740  
Qy 1741 FRVKKINDVOQOASASSAPNKNOLDGKKKPTSPVKPIPONTYRTRVRKNADSKNLAN 1800  
Db 1741 FRVKKINDVOQOASASSAPNKNOLDGKKKPTSPVKPIPONTYRTRVRKNADSKNLAN 1800  
Qy 1801 AERVFSDNKSQKQNLKNNSKDFNDKLPNNEDRVGSGFADFSPHHYTPIEGTPICFGRND 1860  
Db 1801 AERVFSDNKSQKQNLKNNSKDFNDKLPNNEDRVGSGFADFSPHHYTPIEGTPICFGRND 1860  
Qy 1861 SLSLDFDDDDVDLSREKAEIRKAKENKESPAKVTSHTELTSNQQSANKTQATAKQPINR 1920  
Db 1861 SLSLDFDDDDVDLSREKAEIRKAKENKESPAKVTSHTELTSNQQSANKTQATAKQPINR 1920  
Qy 1921 GQPKPILOKQSTFFQSSKDIIPDRGAATDEKLQNFALIENTPVCFSHNSLSLSLSDIDQENN 1980  
Db 1921 GQPKPILOKQSTFFQSSKDIIPDRGAATDEKLQNFALIENTPVCFSHNSLSLSLSDIDQENN 1980  
Qy 1981 NKENEPIKETEPPOSGEPSPKQASGVAPKSFHVEDTPVCFSRNSSLSLSLSDIDEDLLQ 2040  
Db 1981 NKENEPIKETEPPOSGEPSPKQASGVAPKSFHVEDTPVCFSRNSSLSLSLSDIDEDLLQ 2040  
Qy 2041 ECISAMPEKKKPSRLKGDNEKHSFRNMGGLGEBDLTLDLKDIOKPDSEHGLSPDSNF 2100  
Db 2041 ECISAMPEKKKPSRLKGDNEKHSFRNMGGLGEBDLTLDLKDIOKPDSEHGLSPDSNF 2100  
Qy 2101 WKAIQEGANSIVSSLHQAAAAAALSRQASDSDSILSKSGISLGSFPHLTPDQEEKPFT 2160  
Db 2101 WKAIQEGANSIVSSLHQAAAAAALSRQASDSDSILSKSGISLGSFPHLTPDQEEKPFT 2160  
Qy 2161 SNKGPRILLKPEKSTLETFTKTESKGIKGGKVKYKSLITGKVSNSSEI SQMKQPIQAN 2220  
Db 2161 SNKGPRILLKPEKSTLETFTKTESKGIKGGKVKYKSLITGKVSNSSEI SQMKQPIQAN 2220  
Qy 2221 MPSISRGRTMIHIGVRNRSSTSPVSKKGPPLKTPASKSPSEGTATTSPRGAKPSVKS 2280  
Db 2221 MPSISRGRTMIHIGVRNRSSTSPVSKKGPPLKTPASKSPSEGTATTSPRGAKPSVKS 2280  
Qy 2281 ELSVPARQTSQIGSSKAPSRSGRSDTSPSPAQOPLSRPIQSPGRNSISPGRNGISPPN 2340  
Db 2281 ELSVPARQTSQIGSSKAPSRSGRSDTSPSPAQOPLSRPIQSPGRNSISPGRNGISPPN 2340  
Qy 2341 KLSQLPRTSSSTASTKSSGSGKMSYTPSGQMSQOQLTKOTGLSKNASSIPRSESASKG 2400  
Db 2341 KLSQLPRTSSSTASTKSSGSGKMSYTPSGQMSQOQLTKOTGLSKNASSIPRSESASKG 2400



QY 2401 LNMNNGANKKVELSRMSSTKSGESDRERPVLRQSTFIKEAPSPTLRRKLEBSA 2460  
 DB 2401 LNMNNGANKKVELSRMSSTKSGESDRERPVLRQSTFIKEAPSPTLRRKLEBSA 2460  
 QY 2461 SFESLSPSRPASPTRSQAQTPVLSPLPDMSLSTHSSVQAGGWRKLPNLSPTTIEYNDG 2520  
 DB 2461 SFESLSPSRPASPTRSQAQTPVLSPLPDMSLSTHSSVQAGGWRKLPNLSPTTIEYNDG 2520  
 QY 2521 RPAKHDIARSHESPSRLPINRSGTWKREHSHKSSSLPRVSTWRTGSSSILSASSSES 2580  
 DB 2521 RPAKHDIARSHESPSRLPINRSGTWKREHSHKSSSLPRVSTWRTGSSSILSASSSES 2580  
 QY 2581 SEKAKSEDEKHVNSISGTKOSKXENQVSAKGTWRKIKENEFSTNSTSTVSSGATNGAES 2640  
 DB 2581 SEKAKSEDEKHVNSISGTKOSKXENQVSAKGTWRKIKENEFSTNSTSTVSSGATNGAES 2640  
 QY 2641 KTLIYQMAFAVSKTEDVWVRIEDCPINNPRSGRSPGTNTPPVIDSVSEKANPNIKDSKDN 2700  
 DB 2641 KTLIYQMAFAVSKTEDVWVRIEDCPINNPRSGRSPGTNTPPVIDSVSEKANPNIKDSKDN 2700  
 QY 2701 QAKQVNGSVPMRTVGLNRLNSIQVDAPQKGTETKPGQNNPVPVSETNESSIVERT 2760  
 DB 2701 QAKQVNGSVPMRTVGLNRLNSIQVDAPQKGTETKPGQNNPVPVSETNESSIVERT 2760  
 QY 2761 PFSSSSSKHSHSPSGTVAARVTFPNVNPSPRKSADSTSARPSQIPTPVNNTKKRDSKT 2820  
 DB 2761 PFSSSSSKHSHSPSGTVAARVTFPNVNPSPRKSADSTSARPSQIPTPVNNTKKRDSKT 2820  
 QY 2821 DSTESSGTOSPKRHSGLVLTSV 2843  
 DB 2821 DSTESSGTOSPKRHSGLVLTSV 2843

RESULT 2

US-10-392-113-21  
 ; Sequence 21, Application US/10392113  
 ; Publication NO. US20030224993A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Land, Hartmut  
 ; APPLICANT: Deleu, Laurent  
 ; TITLE OF INVENTION: COMPOSITIONS THAT INHIBIT PROLIFERATION  
 ; FILE REFERENCE: 21108.000503  
 ; CURRENT APPLICATION NUMBER: US/10/392,113  
 ; CURRENT FILING DATE: 2003-03-17  
 ; PRIOR APPLICATION NUMBER: 60/365,078  
 ; PRIOR FILING DATE: 2002-03-15  
 ; PRIOR APPLICATION NUMBER: PCI/US01/32127  
 ; PRIOR FILING DATE: 2001-10-15  
 ; PRIOR APPLICATION NUMBER: 60/239,705  
 ; PRIOR FILING DATE: 2000-10-12  
 ; NUMBER OF SEQ ID NOS: 45  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 21  
 ; LENGTH: 2843  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence:/Note =  
 ; OTHER INFORMATION: Synthetic Construct  
 US-10-392-113-21

Query Match 100.0%; Score 14566; DB 12; Length 2843;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2843; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MAASVDQLLKQVEALKVMSNLRQLEDNSNHLTKLETEASNKEVLKQOGSTEDAM 60  
 DB 1 MAASVDQLLKQVEALKVMSNLRQLEDNSNHLTKLETEASNKEVLKQOGSTEDAM 60  
 QY 61 ASSGQIDLLERKELNLDSSNFPVGLKRSQMSLRSGREGSVSSRSGECSPPVPMGSPFR 120  
 DB 61 ASSGQIDLLERKELNLDSSNFPVGLKRSQMSLRSGREGSVSSRSGECSPPVPMGSPFR 120

QY 121 RGFVNGSRESTGYLEELKERSLILLADLKEKEKQWYQAQLQNLTKRIDSLPLTENFSL 180  
 DB 121 RGFVNGSRESTGYLEELKERSLILLADLKEKEKQWYQAQLQNLTKRIDSLPLTENFSL 180  
 QY 181 QDTRQLQLEYEARQIRVAMEBEOIGTCQDMKEKAORRIARIQOIEKDILIRIQLLOQAT 240  
 DB 181 QDTRQLQLEYEARQIRVAMEBEOIGTCQDMKEKAORRIARIQOIEKDILIRIQLLOQAT 240  
 QY 241 EAERSSQNKHETGSHDAERQNEGQVGEINWATSGNGQSTTRMDHETASVLSSSSTHSA 300  
 DB 241 EAERSSQNKHETGSHDAERQNEGQVGEINWATSGNGQSTTRMDHETASVLSSSSTHSA 300  
 QY 301 PRLTSLHGTQKVMYVLSLLMGLTHDKDDMSRLLAMSSQDSCISMRQSGCLPLLIQLL 360  
 DB 301 PRLTSLHGTQKVMYVLSLLMGLTHDKDDMSRLLAMSSQDSCISMRQSGCLPLLIQLL 360  
 QY 361 HGNDKDSVLLGNRSGSEARARASAAALHNIHSDPDDKRGRRIRVHLLEQRAYCETC 420  
 DB 361 HGNDKDSVLLGNRSGSEARARASAAALHNIHSDPDDKRGRRIRVHLLEQRAYCETC 420  
 QY 421 WEMQEAHEPGMDQKPNMPAPVEHQICPAVCVLMKLSFDEEHRHAMNELGGLQAIABLLQ 480  
 DB 421 WEMQEAHEPGMDQKPNMPAPVEHQICPAVCVLMKLSFDEEHRHAMNELGGLQAIABLLQ 480  
 QY 481 VDCMYGLTNDHYSITLRRYAGMALTNLTGPDYVANKATLCMKMGCMRALVAQLKSEEDL 540  
 DB 481 VDCMYGLTNDHYSITLRRYAGMALTNLTGPDYVANKATLCMKMGCMRALVAQLKSEEDL 540  
 QY 541 QQVIASVLRNLSWRADVNSKTLREVGSKVAKMECALEVKEKSTLKSVALNLSAHC 600  
 DB 541 QQVIASVLRNLSWRADVNSKTLREVGSKVAKMECALEVKEKSTLKSVALNLSAHC 600  
 QY 601 ENKADICAVDGAALAFVLGTLYRSQNTLAIIESGGGILRNVSLSIATNEDHRIQILRENN 660  
 DB 601 ENKADICAVDGAALAFVLGTLYRSQNTLAIIESGGGILRNVSLSIATNEDHRIQILRENN 660  
 QY 661 CLOTLLOHLKSHSLTIVSNACGTLNLSARNPKDQEAALWDMGAVSMLKNLIHSHKMIAM 720  
 DB 661 CLOTLLOHLKSHSLTIVSNACGTLNLSARNPKDQEAALWDMGAVSMLKNLIHSHKMIAM 720  
 QY 721 GSAAALRNLMANRPAPKYKDNIMSPGSSLSLHVHRKQKALEAEALDAQHLSETFDNIDNLS 780  
 DB 721 GSAAALRNLMANRPAPKYKDNIMSPGSSLSLHVHRKQKALEAEALDAQHLSETFDNIDNLS 780  
 QY 781 PKASHRSKQKHOSLYGVDYVFDNRHDDNDSNFTGNMTVLSPYLNTTVLPSSSSSRGS 840  
 DB 781 PKASHRSKQKHOSLYGVDYVFDNRHDDNDSNFTGNMTVLSPYLNTTVLPSSSSSRGS 840  
 QY 841 LDSRSEKDRSLERERIGLGNYPATENPGTSSKRGLOISTTAAQIAKYMEEVSAIHTS 900  
 DB 841 LDSRSEKDRSLERERIGLGNYPATENPGTSSKRGLOISTTAAQIAKYMEEVSAIHTS 900  
 QY 901 QEDRSSGTTTELHCVTDERNALRRSSAAHSTHTYFTKSENPNRITCSMPYAKLEYKRSS 960  
 DB 901 QEDRSSGTTTELHCVTDERNALRRSSAAHSTHTYFTKSENPNRITCSMPYAKLEYKRSS 960  
 QY 961 NDSLNSVSSSDGYKRGOMKPSIESYSEDDESFCVGYQPADLAHKIHSANHMDNDGE 1020  
 DB 961 NDSLNSVSSSDGYKRGOMKPSIESYSEDDESFCVGYQPADLAHKIHSANHMDNDGE 1020  
 QY 1021 LDPTINYSKYSDQNLNSGRQSPQNERWAPKHIIIEDEIKQSEQRQSRNQSTTYPVTE 1080  
 DB 1021 LDPTINYSKYSDQNLNSGRQSPQNERWAPKHIIIEDEIKQSEQRQSRNQSTTYPVTE 1080  
 QY 1081 STDDKHLKFPHFQOQECVPSYRSGANGETNRVSGNHGINONVSOGLCOEDDYEDDKP 1140  
 DB 1081 STDDKHLKFPHFQOQECVPSYRSGANGETNRVSGNHGINONVSOGLCOEDDYEDDKP 1140  
 QY 1141 TNXSERYSEEQHEEERPTNYSIKYNEEKRRVDQPIDYSLKYATDIPSSQKQSFSPSKS 1200  
 DB 1141 TNXSERYSEEQHEEERPTNYSIKYNEEKRRVDQPIDYSLKYATDIPSSQKQSFSPSKS 1200

1201 SSGSSKTEHMSSESSENTSTPSSNAKQNLHPSSAQSRSGQPKAATCKVSSINOETIQ 1260  
 1201 SSGSSKTEHMSSESSENTSTPSSNAKQNLHPSSAQSRSGQPKAATCKVSSINOETIQ 1260  
 1261 TYCVEDTPICFSRCSLSLSLSAEDICGNOTTQEADSAANTLQIAIEIKKIGTRSAEDPV 1320  
 1261 TYCVEDTPICFSRCSLSLSLSAEDICGNOTTQEADSAANTLQIAIEIKKIGTRSAEDPV 1320  
 1321 SEVPAVQHRTRKSRRLQGSLSLSAESAARHKAAPFSSGAKSPKSGAQTPKSPPEHYVQET 1380  
 1321 SEVPAVQHRTRKSRRLQGSLSLSAESAARHKAAPFSSGAKSPKSGAQTPKSPPEHYVQET 1380  
 1381 PLMFSTRCTVSSLDSPFSRSTAGSVQGPCGVMVSGIISPSDLPDGPQTMPPSRKTPP 1440  
 1381 PLMFSTRCTVSSLDSPFSRSTAGSVQGPCGVMVSGIISPSDLPDGPQTMPPSRKTPP 1440  
 1441 PPQTAQTKREVKNKAPTAKRESGPKQAANAAVORVQVLPDADTLHLHFAESTPDGF 1500  
 1441 PPQTAQTKREVKNKAPTAKRESGPKQAANAAVORVQVLPDADTLHLHFAESTPDGF 1500  
 1501 SCSSLSLSALSLDEPFIQDVELRIMPVQENDNGNETESQPKESNENQEKAEKTI DSE 1560  
 1501 SCSSLSLSALSLDEPFIQDVELRIMPVQENDNGNETESQPKESNENQEKAEKTI DSE 1560  
 1561 KDLDDDDDDDDIELEBECIISAMPTKSSRKAKPAQATASKLPPPVARKPSQLPVYKLLPS 1620  
 1561 KDLDDDDDDDDIELEBECIISAMPTKSSRKAKPAQATASKLPPPVARKPSQLPVYKLLPS 1620  
 1621 QNRLOPKHVSFTPGDDMPVYCVGFTPIINFSTATSLDITIESPNELAAEGVGGQA 1680  
 1621 QNRLOPKHVSFTPGDDMPVYCVGFTPIINFSTATSLDITIESPNELAAEGVGGQA 1680  
 1681 SGEFEKDDTPTTEGRSTDEAGGKTSSVTIPELDDNKAEGDILAEICINSAMPKGSHPK 1740  
 1681 SGEFEKDDTPTTEGRSTDEAGGKTSSVTIPELDDNKAEGDILAEICINSAMPKGSHPK 1740  
 1741 FRVKKIMDOVQOASASSAPKKNQDCKKKKTPSPVKPIQNTTEYTRVRKNADSKNLN 1800  
 1741 FRVKKIMDOVQOASASSAPKKNQDCKKKKTPSPVKPIQNTTEYTRVRKNADSKNLN 1800  
 1801 AERFSDNKSCKONLKNNSKDFNDKLPNNEDVRGSAFDSPHYTPTEGTPTCYFSRND 1860  
 1801 AERFSDNKSCKONLKNNSKDFNDKLPNNEDVRGSAFDSPHYTPTEGTPTCYFSRND 1860  
 1861 SLSSLDLDDDDVDSREKAEARAKENKESAKVTSTELTNSQOQANKTQAIKQPINR 1920  
 1861 SLSSLDLDDDDVDSREKAEARAKENKESAKVTSTELTNSQOQANKTQAIKQPINR 1920  
 1921 GQPKPILOKQSTPQSSKQIPDRGAATDEKLQFAIENTVCFSHNSLSLSLSDIDQENN 1980  
 1921 GQPKPILOKQSTPQSSKQIPDRGAATDEKLQFAIENTVCFSHNSLSLSLSDIDQENN 1980  
 1981 NKNEPIKETEPDPSQOQEPKQASGAPKSFHVEDTPVCFSRNSLSLSLSDIDDLQ 2040  
 1981 NKNEPIKETEPDPSQOQEPKQASGAPKSFHVEDTPVCFSRNSLSLSLSDIDDLQ 2040  
 2041 ECISAMPKPKKPKRLKDNKESKPRMNGGILGEBDLTLDKIDQRPDSEHGLSPDENFD 2100  
 2041 ECISAMPKPKKPKRLKDNKESKPRMNGGILGEBDLTLDKIDQRPDSEHGLSPDENFD 2100  
 2101 WKALQEGANSIVSLHQAALACLSRAQSDSDSILSKSGISLGSPFHLTPQEEKPPT 2160  
 2101 WKALQEGANSIVSLHQAALACLSRAQSDSDSILSKSGISLGSPFHLTPQEEKPPT 2160  
 2161 SNKGPRILKPKGKSTLETTKTIESKGIKGGKVKYKSLITGKVRNSSEISQWKQPLQAN 2220  
 2161 SNKGPRILKPKGKSTLETTKTIESKGIKGGKVKYKSLITGKVRNSSEISQWKQPLQAN 2220  
 2221 MPSTSRGRMTIHIPGVNRNSSTSPVSKGPPPLKTPASKSPSEGTATTPRGAQKSVKS 2280  
 2221 MPSTSRGRMTIHIPGVNRNSSTSPVSKGPPPLKTPASKSPSEGTATTPRGAQKSVKS 2280  
 2281 ELSPVARTSQIGSSKAPSRSGRSDTSPRPAQQLSRPIQSGRNSISFGRNIGISPPN 2340

1281 ELSPVARTSQIGSSKAPSRSGRSDTSPRPAQQLSRPIQSGRNSISFGRNIGISPPN 2340  
 2341 KLSQLPRTSSPTASTKSSGSKMYTSPGRMQOQNLTKQTGLSKNASSIPRSESASKG 2400  
 2341 KLSQLPRTSSPTASTKSSGSKMYTSPGRMQOQNLTKQTGLSKNASSIPRSESASKG 2400  
 2401 LQNMNGANGANKVELSRMSTKSSGSDRSEBPVLRQSTFIKEAPSPFLRKLSESA 2460  
 2401 LQNMNGANGANKVELSRMSTKSSGSDRSEBPVLRQSTFIKEAPSPFLRKLSESA 2460  
 2461 SFESLSPSSRPASPTRSQAOQTPVLSPLPDMSLSTHSSVQAGGWRKLPPLNLSPTIENDG 2520  
 2461 SFESLSPSSRPASPTRSQAOQTPVLSPLPDMSLSTHSSVQAGGWRKLPPLNLSPTIENDG 2520  
 2521 RPAKEHDIARSHSBSRPLPINRSGTWKREHSHSSSLPASTWRTTSGSSSILSASSES 2580  
 2521 RPAKEHDIARSHSBSRPLPINRSGTWKREHSHSSSLPASTWRTTSGSSSILSASSES 2580  
 2581 SEKAKSEDEKHVNSISGTSKQENQVSAKGTWRKIKENEFSPNTNSTQTSVSSGATGAES 2640  
 2581 SEKAKSEDEKHVNSISGTSKQENQVSAKGTWRKIKENEFSPNTNSTQTSVSSGATGAES 2640  
 2641 KTLIYQMAPAVSKTIEDVWRIEDCPINNPRSGRSPGTNTPPVIDSVSEKANPNIKDSKN 2700  
 2641 KTLIYQMAPAVSKTIEDVWRIEDCPINNPRSGRSPGTNTPPVIDSVSEKANPNIKDSKN 2700  
 2701 QAKQVNGSGVPMRTVGLNRLNSFIOVDAPDQKGTIIPKQNNPVPVSEINNESSIVERT 2760  
 2701 QAKQVNGSGVPMRTVGLNRLNSFIOVDAPDQKGTIIPKQNNPVPVSEINNESSIVERT 2760  
 2761 PFSSSSSKHSSPSTVAARVTPFNYPNPSPKSADSTARSPOIPTPVNNNTKKRUSKT 2820  
 2761 PFSSSSSKHSSPSTVAARVTPFNYPNPSPKSADSTARSPOIPTPVNNNTKKRUSKT 2820  
 2821 DSTESEGTQSPKRGSGVLTSTV 2843  
 2821 DSTESEGTQSPKRGSGVLTSTV 2843

RESULT 3  
 US-10-408-765A-1970  
 ; Sequence 1970, Application US/10408765A  
 ; Publication No. US20040101874A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ghosh, Soumitra S.  
 ; APPLICANT: Fahy, Eoin D.  
 ; APPLICANT: Zhang, Bing  
 ; APPLICANT: Gibson, Bradford W.  
 ; APPLICANT: Taylor, Steven W.  
 ; APPLICANT: Glenn, Gary M.  
 ; APPLICANT: Warnock, Dale E.  
 ; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION  
 ; FILE REFERENCE: 660088.465  
 ; CURRENT APPLICATION NUMBER: US/10/408,765A  
 ; CURRENT FILING DATE: 2003-04-04  
 ; NUMBER OF SEQ ID NOS: 3077  
 ; SOFTWARE: FASTSEQ for Windows Version 4.0  
 ; SEQ ID NO 1970  
 ; LENGTH: 2843  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-408-765A-1970

Query Match 99.9%; Score 14557; DB 16; Length 2843;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2842; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAASVDQLLKQVEALKMENSNLRLQEDNSHLTKLETSANMKVILKQIQGSIEDEAM 60  
 DB 1 MAASVDQLLKQVEALKMENSNLRLQEDNSHLTKLETSANMKVILKQIQGSIEDEAM 60

QY 61 ASSGQIDLLERKELNLDSSNFFPGVKLRKMSLRYSYSGREGSVSRSGECSVPWGMGPPR 120  
DB 61 ASSGQIDLLERKELNLDSSNFFPGVKLRKMSLRYSYSGREGSVSRSGECSVPWGMGPPR 120  
QY 121 RGFVNGSRESTCYLBELEKERSILLADLDKEEKEDWYIAQI ONLTXRIDSJLPTENFSL 180  
DB 121 RGFVNGSRESTCYLBELEKERSILLADLDKEEKEDWYIAQI ONLTXRIDSJLPTENFSL 180  
QY 181 QDMMTRROLEYEARQIRVAMEEQLGTCQDMKRAQRRIARIQQIEKDIRRQLLOSOAT 240  
DB 181 QDMMTRROLEYEARQIRVAMEEQLGTCQDMKRAQRRIARIQQIEKDIRRQLLOSOAT 240  
QY 241 EAERSQNKHETGSHDAERONEGQGVGEINMATSGOGSTTRMDHETASVLSSSSTHSA 300  
DB 241 EAERSQNKHETGSHDAERONEGQGVGEINMATSGOGSTTRMDHETASVLSSSSTHSA 300  
QY 301 PRRLTSHLGTKEYMYVLSLLMLGTHDKDDMGRTLLAMSSSDSCISMROSCLPLLIQLL 360  
DB 301 PRRLTSHLGTKEYMYVLSLLMLGTHDKDDMGRTLLAMSSSDSCISMROSCLPLLIQLL 360  
QY 361 HGNDKDSVLLGNRSRGSKEARAPASAAALHNI THSQPDDKRGREIRVZLHLBQIRAYCETC 420  
DB 361 HGNDKDSVLLGNRSRGSKEARAPASAAALHNI THSQPDDKRGREIRVZLHLBQIRAYCETC 420  
QY 421 NEWQSAHEFGMDODKNPMPAPVEHOICPAVCVLMKLSFDEBHRHAMNELGLOIAELQ 480  
DB 421 NEWQSAHEFGMDODKNPMPAPVEHOICPAVCVLMKLSFDEBHRHAMNELGLOIAELQ 480  
QY 481 VDCMYGLTNDHYSITLARRYAGMALTNLTFDGVANKATLCMKGCMRALVAQLKSESDL 540  
DB 481 VDCMYGLTNDHYSITLARRYAGMALTNLTFDGVANKATLCMKGCMRALVAQLKSESDL 540  
QY 541 QOVIASVLRNLSWRADVNSKTLREVGSVKALMECALEVKESTLKSLSALWLSAHT 600  
DB 541 QOVIASVLRNLSWRADVNSKTLREVGSVKALMECALEVKESTLKSLSALWLSAHT 600  
QY 601 ENKADICAVDGAFLVGLTYRSTOTNLAIIESGGILRVNVSIIATNEBHRQILRENN 660  
DB 601 ENKADICAVDGAFLVGLTYRSTOTNLAIIESGGILRVNVSIIATNEBHRQILRENN 660  
QY 661 CIOTLLOHLKSHSLTIVSNACOTLWNL SARPNKDOEALWDMGAVSMLKNLTHSKHKMIAM 720  
DB 661 CIOTLLOHLKSHSLTIVSNACOTLWNL SARPNKDOEALWDMGAVSMLKNLTHSKHKMIAM 720  
QY 721 GSAALRNLMANRPARYKADANIMSGSSILPSLHVTKOKALFAELDAOHLSETFNIDNLS 780  
DB 721 GSAALRNLMANRPARYKADANIMSGSSILPSLHVTKOKALFAELDAOHLSETFNIDNLS 780  
QY 781 PRASHRSKQRHKQSLYGDVFTDNRHDDNRSDFNTGNMTVLSPLYLNTTVLPSSSSSRGS 840  
DB 781 PRASHRSKQRHKQSLYGDVFTDNRHDDNRSDFNTGNMTVLSPLYLNTTVLPSSSSSRGS 840  
QY 841 LDRSSEKDRSLERERGLGNHYHATENPGTSSFRGLQISTTAAQIAKWEEVSAIHTS 900  
DB 841 LDRSSEKDRSLERERGLGNHYHATENPGTSSFRGLQISTTAAQIAKWEEVSAIHTS 900  
QY 901 QDRSGSGTTELHCVTDENALRRSSAAHTHNTYNTFKSENSNRTCMPYAKLEYKRSS 960  
DB 901 QDRSGSGTTELHCVTDENALRRSSAAHTHNTYNTFKSENSNRTCMPYAKLEYKRSS 960  
QY 961 NDSLNSVSSDGYKRGQMKPISYESVDEBDESKFCYGOYPADLAHKTHSANHMDNDNCE 1020  
DB 961 NDSLNSVSSDGYKRGQMKPISYESVDEBDESKFCYGOYPADLAHKTHSANHMDNDNCE 1020  
QY 1021 LPTPINYSILKYDEQLNSGRQSPONERWARPKHIEDEIKOSEORQSRNOSTTYPVYTE 1080  
DB 1021 LPTPINYSILKYDEQLNSGRQSPONERWARPKHIEDEIKOSEORQSRNOSTTYPVYTE 1080  
QY 1081 STDDKHLKFQPHFGQOECVPSYRSRGANGSETNRYGNSHGINQVYSQSLCQEDDDYDDKP 1140  
DB 1081 STDDKHLKFQPHFGQOECVPSYRSRGANGSETNRYGNSHGINQVYSQSLCQEDDDYDDKP 1140  
QY 1141 TNYSERYSSEEEHREERTNYSIKYNEEKRHVDQPIDYSLKYATDIPSSQKQSFSPKS 1200

DB 1141 TNYSERYSSEEEHREERTNYSIKYNEEKRHVDQPIDYSLKYATDIPSSQKQSFSPKS 1200  
QY 1201 SSGOSSKTEHSSSSSNTTTPSSNAKRONOLHPSSSAQSSRGOPKAATCKVSINQETIQ 1260  
DB 1201 SSGOSSKTEHSSSSSNTTTPSSNAKRONOLHPSSSAQSSRGOPKAATCKVSINQETIQ 1260  
QY 1261 TYCVETDTPICFSCSSLSLSAEDIEIGNQTTQOEADSAANTLQIARIKIGIKTRSAEDPV 1320  
DB 1261 TYCVETDTPICFSCSSLSLSAEDIEIGNQTTQOEADSAANTLQIARIKIGIKTRSAEDPV 1320  
QY 1321 SEVPAYSOHPRTKSSRLQSSLSSESARHKAFFSSGAKSPSKSGAQTPKSPHHVVOET 1380  
DB 1321 SEVPAYSOHPRTKSSRLQSSLSSESARHKAFFSSGAKSPSKSGAQTPKSPHHVVOET 1380  
QY 1381 PLMFSCRTSVSSLDSPESRSIASSVOSEPCSGMVSGIISPSDLPDPSGOTMPPSRSKTTP 1440  
DB 1381 PLMFSCRTSVSSLDSPESRSIASSVOSEPCSGMVSGIISPSDLPDPSGOTMPPSRSKTTP 1440  
QY 1441 PPOQTATQKREVPKNKAPTAERKESGPKQAANVAQVORVOVLDPADTLLHFATESTPDGF 1500  
DB 1441 PPOQTATQKREVPKNKAPTAERKESGPKQAANVAQVORVOVLDPADTLLHFATESTPDGF 1500  
QY 1501 SCSSLSALSLDPEFTQKOVLEIRIMPPVQENDNGNTESEOPKESNENQKEAEKIDSE 1560  
DB 1501 SCSSLSALSLDPEFTQKOVLEIRIMPPVQENDNGNTESEOPKESNENQKEAEKIDSE 1560  
QY 1561 KDLLDSDDDDDIILBECIISAMPTKSSRKAKPAQTASKLPPPVARKPSQLPVYKLLPS 1620  
DB 1561 KDLLDSDDDDDIILBECIISAMPTKSSRKAKPAQTASKLPPPVARKPSQLPVYKLLPS 1620  
QY 1621 QNRLOPKHVSFTPGDDMPREVYCVETGTPINFSTATSLDLTIESPPNELAAGVGVRGAQ 1680  
DB 1621 QNRLOPKHVSFTPGDDMPREVYCVETGTPINFSTATSLDLTIESPPNELAAGVGVRGAQ 1680  
QY 1681 SGFEKEDTPTTEGRSTDEAOGGKTSSVTIPELDDNKAEBGDI LAECINSAMPKGSHP 1740  
DB 1681 SGFEKEDTPTTEGRSTDEAOGGKTSSVTIPELDDNKAEBGDI LAECINSAMPKGSHP 1740  
QY 1741 FRVKIMQVQOASASSAPNKNQLOKGGKXKPSVKPIPONTETRYTRVKNADSKNLN 1800  
DB 1741 FRVKIMQVQOASASSAPNKNQLOKGGKXKPSVKPIPONTETRYTRVKNADSKNLN 1800  
QY 1801 AERFSDNCKSKQNLKNNDKFNNDLPPNEDVRGSAFDPSPHHVTPTEGTYCFSRND 1860  
DB 1801 AERFSDNCKSKQNLKNNDKFNNDLPPNEDVRGSAFDPSPHHVTPTEGTYCFSRND 1860  
QY 1861 SLSSLDPDDDDVLSREKABELRAKENKESEAKVTSHTELTSNQQSANKTQAIKOPINR 1920  
DB 1861 SLSSLDPDDDDVLSREKABELRAKENKESEAKVTSHTELTSNQQSANKTQAIKOPINR 1920  
QY 1921 GQKPILOKQSTTPQSSKDI PDRGAATDEKLOFAIENTVPCFSHNSLSLSDIDQENN 1980  
DB 1921 GQKPILOKQSTTPQSSKDI PDRGAATDEKLOFAIENTVPCFSHNSLSLSDIDQENN 1980  
QY 1981 NKNEPIKETEPDPSQGEPSKQASGYAPKSFHVEDTPVCFSRNSLSLSISIDSEDDLQ 2040  
DB 1981 NKNEPIKETEPDPSQGEPSKQASGYAPKSFHVEDTPVCFSRNSLSLSISIDSEDDLQ 2040  
QY 2041 ECITSSAMPKKKSRUKGDNEKHSPRNMGILGEDITLDIKOTQRPDSEHGLSPDSNFD 2100  
DB 2041 ECITSSAMPKKKSRUKGDNEKHSPRNMGILGEDITLDIKOTQRPDSEHGLSPDSNFD 2100  
QY 2101 WKAIQEGANSIVSSLHQAACCLSRQASSDSISLSKSGISLGSPFHLTPQEEKPFT 2160  
DB 2101 WKAIQEGANSIVSSLHQAACCLSRQASSDSISLSKSGISLGSPFHLTPQEEKPFT 2160  
QY 2161 SNKGPRILKFGEXSTLTETKKIESKGIKGGKVKYKSLITGKVRNSSEISQWKQPLQAN 2220  
DB 2161 SNKGPRILKFGEXSTLTETKKIESKGIKGGKVKYKSLITGKVRNSSEISQWKQPLQAN 2220  
QY 2221 MPSTSRGRTMIHJPGVNSSSSTSPVSKGPPKLPKTPASKPSGQTATTPROAKPSVKS 2280

Db 2221 MPSISRGRTMIHIPVRNSSSTSPVSKGPKLTKPASKSPBEGQTATTSPRGAKPSVKS 2280  
 QY 2281 ELSVARTQIOIGSSKAPERSGRDSTPRPAQPLSRPIQSPGRNSISPGNGISPN 2340  
 Db 2281 ELSVARTQIOIGSSKAPERSGRDSTPRPAQPLSRPIQSPGRNSISPGNGISPN 2340  
 QY 2341 KLSQLPRTSPSTASTKSSGSGKMSYTPGKQMSQQNLTKQTLGSKNASSIPRSESASG 2400  
 Db 2341 KLSQLPRTSPSTASTKSSGSGKMSYTPGKQMSQQNLTKQTLGSKNASSIPRSESASG 2400  
 QY 2401 LQWNGNGANKKVELSRMSTKSSGSDRSPVLRQSTFIKAPSTLRKLEESA 2460  
 Db 2401 LQWNGNGANKKVELSRMSTKSSGSDRSPVLRQSTFIKAPSTLRKLEESA 2460  
 QY 2461 SFESLSPSSRPASPTRQACTPVLSPLPDMSLSTHSSVQAGGWRKLPNLSPTIEYNDG 2520  
 Db 2461 SFESLSPSSRPASPTRQACTPVLSPLPDMSLSTHSSVQAGGWRKLPNLSPTIEYNDG 2520  
 QY 2521 RPAKRHDIAHSHSPRLPINRSGTWKREHSHSSSLPRVSTWRTGSSSSILSASSES 2580  
 Db 2521 RPAKRHDIAHSHSPRLPINRSGTWKREHSHSSSLPRVSTWRTGSSSSILSASSES 2580  
 QY 2581 SEKAKSEDEKHNISIGTKSKENQVSAKGTWKIKENEFSPNSTSQTVSSGATNGAES 2640  
 Db 2581 SEKAKSEDEKHNISIGTKSKENQVSAKGTWKIKENEFSPNSTSQTVSSGATNGAES 2640  
 QY 2641 KTLIYQMAPAVSKTEDVWVRIEDCPINNPRSGSPGTGNTPPVIDSVSEKANPNIKDKN 2700  
 Db 2641 KTLIYQMAPAVSKTEDVWVRIEDCPINNPRSGSPGTGNTPPVIDSVSEKANPNIKDKN 2700  
 QY 2701 OAKQNVGSGVPMRTVGLNRLNSFIOVDAPDQKTEIKPQNNPVVSFTNESSIVERT 2760  
 Db 2701 OAKQNVGSGVPMRTVGLNRLNSFIOVDAPDQKTEIKPQNNPVVSFTNESSIVERT 2760  
 QY 2761 PFSSSSSKHSSPSGTVAARTVPFNPNPSPKSSADSTSAKPSQIPTPVNNNTKKRDSKT 2820  
 Db 2761 PFSSSSSKHSSPSGTVAARTVPFNPNPSPKSSADSTSAKPSQIPTPVNNNTKKRDSKT 2820  
 QY 2821 DSTESSGTQSKRHSGSYLTVS 2843  
 Db 2821 DSTESSGTQSKRHSGSYLTVS 2843

RESULT 4

US-08-681-219-32  
 ; Sequence 32, Application US/08681219  
 ; Publication No. US20020058607A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Takaaki Sato and Junn Yanagisawa  
 ; TITLE OF INVENTION: COMPOUNDS THAT INHIBIT THE INTERACTION BETWEEN  
 ; TITLE OF INVENTION: SIGNAL-TRANSDUCING PROTEINS AND THE GLGF  
 ; TITLE OF INVENTION: (PDZ/DHR) DOMAIN AND USES THEREOF  
 ; NUMBER OF SEQUENCES: 35  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESS: Cooper & Dunham LLP  
 ; STREET: 1185 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: U.S.A.  
 ; ZIP: 10036  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/681,219  
 ; FILING DATE: 22-JUL-1996  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: White, John P  
 ; REGISTRATION NUMBER: 28,678  
 ; REFERENCE/DOCKET NUMBER: 0575/48962/JPW/JKM

TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 278-0400  
 ; TELEFAX: (212) 391-0525  
 ; INFORMATION FOR SEQ ID NO: 32:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 2843 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 ; US-08-681-219-32  
 Query Match 99.8%; Score 14539; DB 8; Length 2843;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 2837; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 MAASDYDQLLKQVAKLWENSNLROELEDNSNHLTKLETSANMKVYLKOLQGSIEDEAM 60  
 Db 1 MAASDYDQLLKQVAKLWENSNLROELEDNSNHLTKLETSANMKVYLKOLQGSIEDEAM 60  
 QY 61 ASSGQIDLLERLKEINLNDSSNFPQVKLRKMSLSRSGSREGSVSSRSGECSPVPMGSPFR 120  
 Db 61 ASSGQIDLLERLKEINLNDSSNFPQVKLRKMSLSRSGSREGSVSSRSGECSPVPMGSPFR 120  
 QY 121 RGFVNGSRESTGYLEELKEKERSILLADLDKEEKEDWYAAQLQNLTKRIDSPLTENFSL 180  
 Db 121 RGFVNGSRESTGYLEELKEKERSILLADLDKEEKEDWYAAQLQNLTKRIDSPLTENFSL 180  
 QY 181 QDTRRQLEYEAKQIRVAMEEQLGTCQDMKRAQRRIARIQIEKDIILRIQLQSQAT 240  
 Db 181 QDTRRQLEYEAKQIRVAMEEQLGTCQDMKRAQRRIARIQIEKDIILRIQLQSQAT 240  
 QY 241 BAERSSQNKHTGSHDAERQNEGQGVGEINMATSGNGQGGSTTRMDHETASVLSSSSTHSA 300  
 Db 241 BAERSSQNKHTGSHDAERQNEGQGVGEINMATSGNGQGGSTTRMDHETASVLSSSSTHSA 300  
 QY 301 PRRLTSHLGTKEVMYSLLSMLGTHDKDMSRTLLAMSSQDSCI SMRQSGCLPLLIQLL 360  
 Db 301 PRRLTSHLGTKEVMYSLLSMLGTHDKDMSRTLLAMSSQDSCI SMRQSGCLPLLIQLL 360  
 QY 361 HGNDKDSVLLGNRSGSKEARARASAAHLNIIHSPDDDKRGRREIRVHLLEQIRAYCETC 420  
 Db 361 HGNDKDSVLLGNRSGSKEARARASAAHLNIIHSPDDDKRGRREIRVHLLEQIRAYCETC 420  
 QY 421 WEQEAHEPGMDQKPNMPAPVEHQICPAVCVLMKLSFDEEHRHANNELGLOQAIALLQ 480  
 Db 421 WEQEAHEPGMDQKPNMPAPVEHQICPAVCVLMKLSFDEEHRHANNELGLOQAIALLQ 480  
 QY 481 VDCMEYGLTNDHYITLRYAGMALTNLTGADVANKATLCMKGCMRALVAQLKSEEDL 540  
 Db 481 VDCMEYGLTNDHYITLRYAGMALTNLTGADVANKATLCMKGCMRALVAQLKSEEDL 540  
 QY 541 QQVIAVLRNLNSWADVNSKKTIREVGSVKALMECALEVKKESTLKSIVLSALNLSAHT 600  
 Db 541 QQVIAVLRNLNSWADVNSKKTIREVGSVKALMECALEVKKESTLKSIVLSALNLSAHT 600  
 QY 601 ENKADICAVDQALAFVLGTITLTYRSQTNTLAIIESGGGILRNVSLSIATNEDHRLRENN 660  
 Db 601 ENKADICAVDQALAFVLGTITLTYRSQTNTLAIIESGGGILRNVSLSIATNEDHRLRENN 660  
 QY 661 CLQTLQHLKSHSLTIIVSNACGTLWNLISARNPKDQEBALWDMGAVSMKLNLIHSHKMIAM 720  
 Db 661 CLQTLQHLKSHSLTIIVSNACGTLWNLISARNPKDQEBALWDMGAVSMKLNLIHSHKMIAM 720  
 QY 721 GSAALRNLMANRPKADANIMSPGSSLSPLSHVRKOKALEAEALDAHLSETFDNIDLS 780  
 Db 721 GSAALRNLMANRPKADANIMSPGSSLSPLSHVRKOKALEAEALDAHLSETFDNIDLS 780  
 QY 781 PKASHRSKQRHKQSLYGDYVDFVTNRHDDNRSDNPNCTGMVTLSPYLNNTVLPSSSSSRGS 840  
 Db 781 PKASHRSKQRHKQSLYGDYVDFVTNRHDDNRSDNPNCTGMVTLSPYLNNTVLPSSSSSRGS 840  
 QY 841 LDSSRSKORSLEERERGIGLGNVHPATENPGTSSKRGLOISTTAAQIAKVMEVSAHTS 900

Db 841 LDSSRSKDRSLERERGI GLGNYPATENPGTSSKRGLOISTTAAQIAKWEEVSAIHTS 900  
QY 901 QEDRSSGTTTLHCYTDERNALRRSSAAHTSNFTKSENSNRTCSMPYAKLEYKSS 960  
Db 901 QEDRSSGTTTLHCYTDERNALRRSSAAHTSNFTKSENSNRTCSMPYAKLEYKSS 960  
QY 961 NDLSNSVSSDGYGKRGQKPSIESYSDDESCKFCYGYQYPADLAHKIHSANHMDNDGE 1020  
Db 961 NDLSNSVSSDGYGKRGQKPSIESYSDDESCKFCYGYQYPADLAHKIHSANHMDNDGE 1020  
QY 1021 LDTPINYLKYSDDEOLNRSRQSPQSNVERWAPKPHIEDEIKQSEORQSRNQSTTTPVYTE 1080  
Db 1021 LDTPINYLKYSDDEOLNRSRQSPQSNVERWAPKPHIEDEIKQSEORQSRNQSTTTPVYTE 1080  
QY 1081 STDDKHLKFPQHFQGOECVPSVRSRGANGSETNRVGSNHGINQVNSQSLCOEDDEYDDKP 1140  
Db 1081 STDDKHLKFPQHFQGOECVPSVRSRGANGSETNRVGSNHGINQVNSQSLCOEDDEYDDKP 1140  
QY 1141 TNYISRYSEBEOHEBERPTNYSIKYNEEKRHVDQPIDYSILKYATDIPSSQKQSFSPKS 1200  
Db 1141 TNYISRYSEBEOHEBERPTNYSIKYNEEKRHVDQPIDYSILKYATDIPSSQKQSFSPKS 1200  
QY 1201 SSGQSKTEHMSSENSTSPSSNAKRONLHPSSAQSRSGOPQKAAATCKVSSINOETIQ 1260  
Db 1201 SSGQSKTEHMSSENSTSPSSNAKRONLHPSSAQSRSGOPQKAAATCKVSSINOETIQ 1260  
QY 1261 TYCVEDTPICFRCSLSLSAEDIEICGNQTTQPADSANTLOIAIEIKEIGTRSAEDPV 1320  
Db 1261 TYCVEDTPICFRCSLSLSAEDIEICGNQTTQPADSANTLOIAIEIKEIGTRSAEDPV 1320  
QY 1321 SEVPAVSQHPRTKSRLOQSSLSSESARHKAVEFSSGAKSPKSCAQTPKPPPEHYVOET 1380  
Db 1321 SEVPAVSQHPRTKSRLOQSSLSSESARHKAVEFSSGAKSPKSCAQTPKPPPEHYVOET 1380  
QY 1381 PLMFRCSTSVSLDSFESRSIASSVQSPCSGMVSGIISPDLDPSPQGTWPPRSKTKTP 1440  
Db 1381 PLMFRCSTSVSLDSFESRSIASSVQSPCSGMVSGIISPDLDPSPQGTWPPRSKTKTP 1440  
QY 1441 PPQTAQTKREVPKNAKPTAEKRESGPKQAAVNAVQVQLPDADTLLHFATSTPDGF 1500  
Db 1441 PPQTAQTKREVPKNAKPTAEKRESGPKQAAVNAVQVQLPDADTLLHFATSTPDGF 1500  
QY 1501 SCSSLSALSLEPFIQKDVLRIMPVQENDNGNETSEQPKENQENQEAETIDSE 1560  
Db 1501 SCSSLSALSLEPFIQKDVLRIMPVQENDNGNETSEQPKENQENQEAETIDSE 1560  
QY 1561 KOLLDDSDDDIIEILEECIISAMPTKSRKAKKPAQTASKLPPPVARKPSQLPVYKLLPS 1620  
Db 1561 KOLLDDSDDDIIEILEECIISAMPTKSRKAKKPAQTASKLPPPVARKPSQLPVYKLLPS 1620  
QY 1621 QNRLOPKHVSFTPGDDMPRVYCVETGTPINFSTATSLSLTIESPNELAAEGVGRGAQ 1680  
Db 1621 QNRLOPKHVSFTPGDDMPRVYCVETGTPINFSTATSLSLTIESPNELAAEGVGRGAQ 1680  
QY 1681 SGFEFKRDTIPTEGRSTDEAQGGKTSVTPPELDONKAEEDGDLAECINSAMPKGSKHP 1740  
Db 1681 SGFEFKRDTIPTEGRSTDEAQGGKTSVTPPELDONKAEEDGDLAECINSAMPKGSKHP 1740  
QY 1741 FRVKTMDOVQQAASASAPNKNQLDGKKKTPSPVKPIPQNTETATRVKRNKADSKNIN 1800  
Db 1741 FRVKTMDOVQQAASASAPNKNQLDGKKKTPSPVKPIPQNTETATRVKRNKADSKNIN 1800  
QY 1801 AERVFSDNKSCKONLKNNSKDFNDKLPNNEDRVGSPAFDGFPHYTPTEGTPYCFSRND 1860  
Db 1801 AERVFSDNKSCKONLKNNSKDFNDKLPNNEDRVGSPAFDGFPHYTPTEGTPYCFSRND 1860  
QY 1861 SLSSLDFFDDDDVLSREKAEKAKENKSEAKVTSHTELTNSQOSANKTOIAKQPINR 1920  
Db 1861 SLSSLDFFDDDDVLSREKAEKAKENKSEAKVTSHTELTNSQOSANKTOIAKQPINR 1920  
QY 1921 GQPKPILOKSTPPQSKDIPDRGAATDEKLNQFAIENTPVCFSHNSLSLSDIDQENN 1980

Db 1921 GQPKPILOKSTPPQSKDIPDRGAATDEKLNQFAIENTPVCFSHNSLSLSDIDQENN 1980  
QY 1981 KNENEPIKETEPDPSOGEPSPKQASGYAPKSPHVEDTPVCFGRNSLSLSLSDIDQENN 2040  
Db 1981 KNENEPIKETEPDPSOGEPSPKQASGYAPKSPHVEDTPVCFGRNSLSLSLSDIDQENN 2040  
QY 2041 ECTISSAMPKKKPSRLKGDNEKHSRNMGGIIGEDLTLDLKOIQRDPSHGLSPDSENF 2100  
Db 2041 ECTISSAMPKKKPSRLKGDNEKHSRNMGGIIGEDLTLDLKOIQRDPSHGLSPDSENF 2100  
QY 2101 WKAIQEGANSIVSSLSHQAACLSRQASDSDSILSLKSGISLGSPPFHLTPQBEKPF 2160  
Db 2101 WKAIQEGANSIVSSLSHQAACLSRQASDSDSILSLKSGISLGSPPFHLTPQBEKPF 2160  
QY 2161 SNKGPRILKPEKSTLTETKKIESESKIGKGVKYSKLITGKVRNSSEISGQWKQLOAN 2220  
Db 2161 SNKGPRILKPEKSTLTETKKIESESKIGKGVKYSKLITGKVRNSSEISGQWKQLOAN 2220  
QY 2221 MPGISRGRTMIHPTGVNNSSTSPVSKGPPKLTTPASKSPSEGTATTPRCAKPSVKS 2280  
Db 2221 MPGISRGRTMIHPTGVNNSSTSPVSKGPPKLTTPASKSPSEGTATTPRCAKPSVKS 2280  
QY 2281 ELSPVARQTSQIGGSSKAPSRGSRDSTPSRPAQQLSRPQPSGRNSISPGRNGISPPN 2340  
Db 2281 ELSPVARQTSQIGGSSKAPSRGSRDSTPSRPAQQLSRPQPSGRNSISPGRNGISPPN 2340  
QY 2341 KLSQLPRTSPSTASTKSSGSGKSVYTSRGRMSQOONLTKOTGLSKYNASSIPSEASGK 2400  
Db 2341 KLSQLPRTSPSTASTKSSGSGKSVYTSRGRMSQOONLTKOTGLSKYNASSIPSEASGK 2400  
QY 2401 LNMWNGNGANKVELSRMSSTKSSGSESRRSRPVLVROSTFIKEAPSTLARKLEESA 2460  
Db 2401 LNMWNGNGANKVELSRMSSTKSSGSESRRSRPVLVROSTFIKEAPSTLARKLEESA 2460  
QY 2461 SPESLSPSPSPASTRSQAOTPVLSPLPMSLSTHSSVOAGGWRKLPPLNSTIETNDG 2520  
Db 2461 SPESLSPSPSPASTRSQAOTPVLSPLPMSLSTHSSVOAGGWRKLPPLNSTIETNDG 2520  
QY 2521 RPAKGRHDIASHSPSRLPINRSGTWKREHSHSSLPVSTWRTTSGSSSILSASSES 2580  
Db 2521 RPAKGRHDIASHSPSRLPINRSGTWKREHSHSSLPVSTWRTTSGSSSILSASSES 2580  
QY 2581 SEKAKSEDEKHVNSISGTSKQENQVSAKWTWAKIKENEFSPNTSQTQVSSGATNGAES 2640  
Db 2581 SEKAKSEDEKHVNSISGTSKQENQVSAKWTWAKIKENEFSPNTSQTQVSSGATNGAES 2640  
QY 2641 KTLIYQAPAVSKTEDVWVRIEDCPINNPRSGSPCTGTPPVIDSYSEKANPNIKSDKN 2700  
Db 2641 KTLIYQAPAVSKTEDVWVRIEDCPINNPRSGSPCTGTPPVIDSYSEKANPNIKSDKN 2700  
QY 2701 QAKQNVNGSVPMRTVGLNRLNSFIQVDAPDQKGTBIKPGQNNPVPVSETNESSIVERT 2760  
Db 2701 QAKQNVNGSVPMRTVGLNRLNSFIQVDAPDQKGTBIKPGQNNPVPVSETNESSIVERT 2760  
QY 2761 PFSSSSSKHSSPSGTVAARVTPNPNPSPRKSSADSTSRPSQIPTPVNNNTKKRDSKT 2820  
Db 2761 PFSSSSSKHSSPSGTVAARVTPNPNPSPRKSSADSTSRPSQIPTPVNNNTKKRDSKT 2820  
QY 2821 DSTESSGTQSPKRHSGSYLVTSV 2843  
Db 2821 DSTESSGTQSPKRHSGSYLVTSV 2843

RESULT 5  
US-09-230-111C-30  
; Sequence 30, Application US/09230111C  
; Publication No. US20030203414A1  
; GENERAL INFORMATION:  
; APPLICANT: Sato, Taka-Aki  
; APPLICANT: Yanagisawa, Junn  
; TITLE OF INVENTION: COMPOUNDS THAT INHIBIT INTERACTION BETWEEN  
; TITLE OF INVENTION: SIGNAL-TRANSDUCING PROTEINS AND THE GLGF (PDZ/DHR)  
; TITLE OF INVENTION: DOMAIN AND USES THEREOF

FILE REFERENCE: 48962-A-PCT-US  
 CURRENT APPLICATION NUMBER: US/09/230.111C  
 CURRENT FILING DATE: 1999-05-17  
 NUMBER OF SEQ ID NOS: 33  
 SOFTWARE: PatentIn ver. 2.1  
 SEQ ID NO 30  
 LENGTH: 2843  
 TYPE: PRT  
 ORGANISM: human  
 US-09-230-111C-30

Query Match 99.8%; Score 14539; DB 11; Length 2843;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 2837; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY	1	MAAASYDQLLKQVEALXWENLNRLQELNDSNHLTKLEASNNKEVYLKOLQSGSIEDAM	60	841	LDSSRSKDRSLERERIGLGNYPATENPGTSSKRGQLISTTAAQIAKVAKEEVSIAHTS	900
DB	1	MAAASYDQLLKQVEALXWENLNRLQELNDSNHLTKLEASNNKEVYLKOLQSGSIEDAM	60	901	QEDRSSGTTTELHCVTDERNALRRSSAAHSHNTNTFTKSENSNRTCSNMPYAKLEYKRSS	960
QY	61	ASSGQIDLLERLKEINLNDSSNFPQGVKLRSKMSLRSYSGREGSVSSRSGECPVPMGSPFR	120	901	QEDRSSGTTTELHCVTDERNALRRSSAAHSHNTNTFTKSENSNRTCSNMPYAKLEYKRSS	960
DB	61	ASSGQIDLLERLKEINLNDSSNFPQGVKLRSKMSLRSYSGREGSVSSRSGECPVPMGSPFR	120	961	NDLSNVSSSDGYGKRGQWKPSIESYSDDDSKFQSYGQYADLAHAKTHSANHMDNDGE	1020
QY	121	RGFVNGRESGYLEELKERSLLADLDKEEKEKDYAQLQNLTKRIDSLPLTENFSL	180	961	NDLSNVSSSDGYGKRGQWKPSIESYSDDDSKFQSYGQYADLAHAKTHSANHMDNDGE	1020
DB	121	RGFVNGRESGYLEELKERSLLADLDKEEKEKDYAQLQNLTKRIDSLPLTENFSL	180	1021	LDTPINYSLKYSDEQLNSGRQSPQNERWARPKHIIIEDEIKQSEORQSRNQSTTYPVYTE	1080
QY	181	QTDWTRQLEYEARQIRVAMEEQLGTQDMKKAQRRIARIQIEKDILIRQLLOSQAT	240	1021	LDTPINYSLKYSDEQLNSGRQSPQNERWARPKHIIIEDEIKQSEORQSRNQSTTYPVYTE	1080
DB	181	QTDWTRQLEYEARQIRVAMEEQLGTQDMKKAQRRIARIQIEKDILIRQLLOSQAT	240	1081	STDDKHLKFPQHFQGCVCVSPYRSGANGSETNRVGSNHGINQNVOSLCCEDDYEDDKP	1140
QY	241	EAERSSQNKHETGSHDAERNEGQGVGEINMATSGNGQGGTTRMDHETASVLSSSSTHSA	300	1081	STDDKHLKFPQHFQGCVCVSPYRSGANGSETNRVGSNHGINQNVOSLCCEDDYEDDKP	1140
DB	241	EAERSSQNKHETGSHDAERNEGQGVGEINMATSGNGQGGTTRMDHETASVLSSSSTHSA	300	1141	TNYSERYSEEEQHEEERPTNYSIKYNEEKRHVDQPDYSLKYATDIPSSQKQSFSPKS	1200
QY	301	PRRLTSHLGTKEVMSYLLSGLTHDKDMSRTLLAMSSQSDSCINMRQSGCLPLLIQLL	360	1141	TNYSERYSEEEQHEEERPTNYSIKYNEEKRHVDQPDYSLKYATDIPSSQKQSFSPKS	1200
DB	301	PRRLTSHLGTKEVMSYLLSGLTHDKDMSRTLLAMSSQSDSCINMRQSGCLPLLIQLL	360	1201	SSGQSKTEHMSSESTSTPSSNAKRONQLHPSAQSRSQPOKAATCKVSSINQETIQ	1260
QY	361	HGNDKDSVLLGNSRGSKEARASAAALHNIHSQPDQKRGREIRVLHLLEQIRAYCETC	420	1201	SSGQSKTEHMSSESTSTPSSNAKRONQLHPSAQSRSQPOKAATCKVSSINQETIQ	1260
DB	361	HGNDKDSVLLGNSRGSKEARASAAALHNIHSQPDQKRGREIRVLHLLEQIRAYCETC	420	1261	TYCVEDTPICFRSCSSLSLSSAEDEIGCNOTTOEADSANTLQIAEIKIGTSAEDPV	1320
QY	421	WWEQEAHEPGMDQKNPMPAPVEHOICPACVLMKLSFDEEHRHANNELGGUQAIABELQ	480	1261	TYCVEDTPICFRSCSSLSLSSAEDEIGCNOTTOEADSANTLQIAEIKIGTSAEDPV	1320
DB	421	WWEQEAHEPGMDQKNPMPAPVEHOICPACVLMKLSFDEEHRHANNELGGUQAIABELQ	480	1321	SEVPASVQHPRTKSRLOGSSLSSESAPHKAVFESSGAKSPSKGAGTTPKPPPHYVQET	1380
QY	481	VDCEMYGLTNDHYSITLRRYAGMALTNLTFGDVANKATLCSMKGCMRALVAQLKSEEDL	540	1321	SEVPASVQHPRTKSRLOGSSLSSESAPHKAVFESSGAKSPSKGAGTTPKPPPHYVQET	1380
DB	481	VDCEMYGLTNDHYSITLRRYAGMALTNLTFGDVANKATLCSMKGCMRALVAQLKSEEDL	540	1381	PLMFRCCTSVSLDSFESRSIASSVQSEPCSGMVGIIISPSDLPSDQTPMPPSRSKTTP	1440
QY	541	QOVIASVLRNLWRADVNSKTLREVGSVKALMECALEVVKESTLKSVALNLSAHT	600	1381	PLMFRCCTSVSLDSFESRSIASSVQSEPCSGMVGIIISPSDLPSDQTPMPPSRSKTTP	1440
DB	541	QOVIASVLRNLWRADVNSKTLREVGSVKALMECALEVVKESTLKSVALNLSAHT	600	1441	PPPTAQTKREVVKPKAKAPTAEKRESGPKQAANAQVQVLPDADTLHLHPATESTPDGF	1500
QY	601	ENKADI CAVDGALAFVGLTYRSQTNLTALIESGGGILRNVSLLIATNEDHRLRENN	660	1441	PPPTAQTKREVVKPKAKAPTAEKRESGPKQAANAQVQVLPDADTLHLHPATESTPDGF	1500
DB	601	ENKADI CAVDGALAFVGLTYRSQTNLTALIESGGGILRNVSLLIATNEDHRLRENN	660	1501	SCSSLSLSLSDLEPFIQKVELRIMPPVQENDNGNETESQPKESNENQKEAEKTTDSE	1560
QY	661	CIQTLLQHLKSHSLTIVSNACGLTWNLSARNPKDQALWDMGVSMKLNLIHSKHQMIAM	720	1501	SCSSLSLSLSDLEPFIQKVELRIMPPVQENDNGNETESQPKESNENQKEAEKTTDSE	1560
DB	661	CIQTLLQHLKSHSLTIVSNACGLTWNLSARNPKDQALWDMGVSMKLNLIHSKHQMIAM	720	1561	KDLLDDSDDDDIIELEECIISAMPTKSSRAKKAQATASKLPPPVARKPSQLPVYKLLPS	1620
QY	721	GSAAALRNLMANRPKADANIMSGSSLPVLRVQKQKALEAELDAQHLSSETFDNIDLS	780	1561	KDLLDDSDDDDIIELEECIISAMPTKSSRAKKAQATASKLPPPVARKPSQLPVYKLLPS	1620
DB	721	GSAAALRNLMANRPKADANIMSGSSLPVLRVQKQKALEAELDAQHLSSETFDNIDLS	780	1621	QNRLOPKHVSFTGDDMPRVYCVETGPINFSTATSLDITIESPPNELAAGEVGRGAQ	1680
QY	781	PKASHRSKQBKHQSLYGVYVDFDTRNHDNRDNFNFTGNMTVLSPLYNTTLVPSSSSSRGS	840	1621	QNRLOPKHVSFTGDDMPRVYCVETGPINFSTATSLDITIESPPNELAAGEVGRGAQ	1680
DB	781	PKASHRSKQBKHQSLYGVYVDFDTRNHDNRDNFNFTGNMTVLSPLYNTTLVPSSSSSRGS	840	1681	SGETEKRTDITPTEGRSTDEAOGGKTSSTVITPELDDNKAEEGDIIAECINSAMPKGGKSHK	1740
QY	841	LDSSRSPKDSIERRPCTICNVHPATENPCTSSKRCIOTSTTAAQIAKVAKEEVSIAHTS	900	1681	SGETEKRTDITPTEGRSTDEAOGGKTSSTVITPELDDNKAEEGDIIAECINSAMPKGGKSHK	1740



1981 NKENPPIKETPPDQGGPSPKQASGYAPKPFHVEDTPVCFSRNSSLSSLSIDSEDLQ 2040  
1981 NKENPPIKETPPDQGGPSPKQASGYAPKPFHVEDTPVCFSRNSSLSSLSIDSEDLQ 2040  
2041 ECISAMPKKKPRLKGDNEKHSRNNGGILGEDLTLDLKDIOQPDSEHGLSPDSENF 2100  
2041 ECISAMPKKKPRLKGDNEKHSRNNGGILGEDLTLDLKDIOQPDSEHGLSPDSENF 2100  
2101 WKAIQEGANSIVSSLHQAAAACLSRQASSDSDLSLSKSGISLGSPPHLPDQEEPPFT 2160  
2101 WKAIQEGANSIVSSLHQAAAACLSRQASSDSDLSLSKSGISLGSPPHLPDQEEPPFT 2160  
2161 SNKGRPIILKPKGEKSTLETKKIESEKGIKGGKVKYKSLITKVRNSNBSISQOMKQPLQAN 2220  
2161 SNKGRPIILKPKGEKSTLETKKIESEKGIKGGKVKYKSLITKVRNSNBSISQOMKQPLQAN 2220  
2221 MPSISGRGRTMIHIPGVRNNSSSSTSPVSKGPKLPKTPASKSPSEGOATTTSPRGAKPSVKS 2280  
2221 MPSISGRGRTMIHIPGVRNNSSSSTSPVSKGPKLPKTPASKSPSEGOATTTSPRGAKPSVKS 2280  
2281 ELSPVARTQSIGGSKAPSRGSDSTPSRPAQOPLSRPQSPGRNBSISQOMKQPLQAN 2340  
2281 ELSPVARTQSIGGSKAPSRGSDSTPSRPAQOPLSRPQSPGRNBSISQOMKQPLQAN 2340  
2341 KLSQLPRTSSPSTATSKSGGKMSYTPSGQMSQONLTQGLSKNASSIPRSESASKG 2400  
2341 KLSQLPRTSSPSTATSKSGGKMSYTPSGQMSQONLTQGLSKNASSIPRSESASKG 2400  
2401 LNQMNGNGANKKVELSRMSTKSSGSDSDRSEPVLRQSTFIKEAPSPTLRRKLEESA 2460  
2401 LNQMNGNGANKKVELSRMSTKSSGSDSDRSEPVLRQSTFIKEAPSPTLRRKLEESA 2460  
2461 SPESLPSRSPASPTSCAQTPLSPSLPDMSLSTHSSVOAGGWRKLPNLSPTIEYNDG 2520  
2461 SPESLPSRSPASPTSCAQTPLSPSLPDMSLSTHSSVOAGGWRKLPNLSPTIEYNDG 2520  
2521 RPAKHDDIARSHSESPSRIPINRSGTWKREHSHSSSIPRVSTWRTGSSSSSILSASSES 2580  
2521 RPAKHDDIARSHSESPSRIPINRSGTWKREHSHSSSIPRVSTWRTGSSSSSILSASSES 2580  
2581 SEKAKSEDEKHVNSISGTKQKXENQVSAKGTWRKIKENEFPSTNSTQTSVSSGATNGAES 2640  
2581 SEKAKSEDEKHVNSISGTKQKXENQVSAKGTWRKIKENEFPSTNSTQTSVSSGATNGAES 2640  
2641 KTLIIQMAPAVSKTDDVVRIEDCPINNPRSGRSPGTNTPPVIDSVSEKAMPNIKSDKN 2700  
2641 KTLIIQMAPAVSKTDDVVRIEDCPINNPRSGRSPGTNTPPVIDSVSEKAMPNIKSDKN 2700  
2701 QAKQNVGSGVPMRTVGLNRLNSFIQVADPQKGTETKPGQNNPVPVSETNESSIVERT 2760  
2701 QAKQNVGSGVPMRTVGLNRLNSFIQVADPQKGTETKPGQNNPVPVSETNESSIVERT 2760  
2761 PFSSSSSKHSGSPGTVAARVTPFNYPNPSPRKSSADSTSRPSQIPTPVNNTKKRDSKT 2820  
2761 PFSSSSSKHSGSPGTVAARVTPFNYPNPSPRKSSADSTSRPSQIPTPVNNTKKRDSKT 2820  
2821 DSTESSGTSQPRHSGSYLVTSV 2843  
2821 DSTESSGTSQPRHSGSYLVTSV 2843

## RESULT 6

US-10-092-138-30

; Sequence 30, Application US/10092138

; Publication No. US20030170723A1

; GENERAL INFORMATION:

; APPLICANT: Sato, Taka-Aki

; TITLE OF INVENTION: METHOD OF PREPARING A PROTEIN ARRAY BASED ON

; TITLE OF INVENTION: BIOCHEMICAL PROTEIN-PROTEIN INTERACTION

; FILE REFERENCE: 65823/JPW/PT

; CURRENT APPLICATION NUMBER: US/10/092,138

; CURRENT FILING DATE: 2002-09-06

; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 30  
; LENGTH: 2843  
; TYPE: PRT  
; ORGANISM: human  
US-10-092-138-30

## Query Match

99.8%; Score 14539; DB 14; Length 2843;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 2837; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MAASVDQLLKQVEALKMENSILRQLEDNSNHLTKLETEASNKMKVLKOLQGSIDEAM 60  
Db 1 MAASVDQLLKQVEALKMENSILRQLEDNSNHLTKLETEASNKMKVLKOLQGSIDEAM 60  
Qy 61 ASSGQIDLLERLKLNDSSNFFGVKLRKMSLRYSRGSGSVSSRGSGECPVPMGSPFR 120  
Db 61 ASSGQIDLLERLKLNDSSNFFGVKLRKMSLRYSRGSGSVSSRGSGECPVPMGSPFR 120  
Qy 121 RGFVNGSRESTGYLEBELEKERSILLADLKEEKEKDYIAQLQNLTKRIDLPLETFNSL 180  
Db 121 RGFVNGSRESTGYLEBELEKERSILLADLKEEKEKDYIAQLQNLTKRIDLPLETFNSL 180  
Qy 181 QTDWTRQLEYEARQIRVAMEEQLGTCQDMEXKAQRRIARIQOIEKDILIRIOLLOQAT 240  
Db 181 QTDWTRQLEYEARQIRVAMEEQLGTCQDMEXKAQRRIARIQOIEKDILIRIOLLOQAT 240  
Qy 241 EAERSSONKHETGSHDAERQNEGGVGEINMATSNGNGQSTTRMDHETASVLSSTHSA 300  
Db 241 EAERSSONKHETGSHDAERQNEGGVGEINMATSNGNGQSTTRMDHETASVLSSTHSA 300  
Qy 301 PRLTSHLGTKEWVYSLLSMLGTHDKDDMSRTLLAMSSQSDCISMRQSGCLPLLIQLL 360  
Db 301 PRLTSHLGTKEWVYSLLSMLGTHDKDDMSRTLLAMSSQSDCISMRQSGCLPLLIQLL 360  
Qy 361 HGNDKDSVLLGNRSGSKEARASAAALHNIHISQDPDKRGRREIRVHLHLEQIRAYCETC 420  
Db 361 HGNDKDSVLLGNRSGSKEARASAAALHNIHISQDPDKRGRREIRVHLHLEQIRAYCETC 420  
Qy 421 WEWQEAHEPGMDQKNMPAPAVEHQICPAVCVLMKLSFDEEHHAMNGLGQIAELQL 480  
Db 421 WEWQEAHEPGMDQKNMPAPAVEHQICPAVCVLMKLSFDEEHHAMNGLGQIAELQL 480  
Qy 481 VDCMYGLTNDHYSITLRRYAGMALTNLTFDGVANKATLCSMGKCMRVAQKLSSEDL 540  
Db 481 VDCMYGLTNDHYSITLRRYAGMALTNLTFDGVANKATLCSMGKCMRVAQKLSSEDL 540  
Qy 541 QQVITASVLRNLSWRADVNSKTLREVGSKVAMECALFVKKESTLKSVALNLSAHCT 600  
Db 541 QQVITASVLRNLSWRADVNSKTLREVGSKVAMECALFVKKESTLKSVALNLSAHCT 600  
Qy 601 ENKADICAVDGAFLAVGLTYRSQNTLAIIESGGILRNVSLLIATNEDHQILRENN 660  
Db 601 ENKADICAVDGAFLAVGLTYRSQNTLAIIESGGILRNVSLLIATNEDHQILRENN 660  
Qy 661 CLOTLLOHLSHSLTIVSNACGTLWNLSARNPKDQBALMDMGAVSMLKNLIHSHKXMIAM 720  
Db 661 CLOTLLOHLSHSLTIVSNACGTLWNLSARNPKDQBALMDMGAVSMLKNLIHSHKXMIAM 720  
Qy 721 GSAALLENIMANEPKVKDANIMSPGSSLPSEVRKQKALEAELDAQHSETDNDNIS 780  
Db 721 GSAALLENIMANEPKVKDANIMSPGSSLPSEVRKQKALEAELDAQHSETDNDNIS 780  
Qy 781 PKASHRSKQKHOSLYGDDYVDFDNRHDDNRSDFNFTGNMTVLSPLYLNTTLPSSSSSRGS 840  
Db 781 PKASHRSKQKHOSLYGDDYVDFDNRHDDNRSDFNFTGNMTVLSPLYLNTTLPSSSSSRGS 840  
Qy 841 LDSRSSEKRSLEIRERIGLGNTHPATENPGTSSKGLQISTTAAQIAKMEVSAIHTS 900  
Db 841 LDSRSSEKRSLEIRERIGLGNTHPATENPGTSSKGLQISTTAAQIAKMEVSAIHTS 900  
Qy 901 QEDRSSGSGTTELHCVTDERNALRRSSAAATHSTNTYNTFKSENSNRITCSMPYAKLEYKRSS 960



Db 901 QEDRSSGTTTLCVTDERNALRRSSAAHSHNTYNTFTKSENRR:CSMPYAKLEYKRS 960  
Qy 961 NDSLSNVSSDGYGKRCQMKPS:IESXSEDDSEKFCYGYQYFADLAHKIHSANHMDNDGE 1020  
Db 961 NDSLSNVSSDGYGKRCQMKPS:IESXSEDDSEKFCYGYQYFADLAHKIHSANHMDNDGE 1020  
Qy 1021 LDPINTYSLKYSEBOLNCSQSPQSONERWAPXHI:IEDEIKQSEQRQSNQSTTYPVYTE 1080  
Db 1021 LDPINTYSLKYSEBOLNCSQSPQSONERWAPXHI:IEDEIKQSEQRQSNQSTTYPVYTE 1080  
Qy 1081 STDDKH:KFQPHFGQCBVPSYRSGANGSETNRVGSNHNINQVQSILCQEDDYEDDKP 1140  
Db 1081 STDDKH:KFQPHFGQCBVPSYRSGANGSETNRVGSNHNINQVQSILCQEDDYEDDKP 1140  
Qy 1141 TNSYRYSBEOHEBERPTNY:IKNEBKRVHVDQIDYSLKYATDIPSSQKQSPFSKS 1200  
Db 1141 TNSYRYSBEOHEBERPTNY:IKNEBKRVHVDQIDYSLKYATDIPSSQKQSPFSKS 1200  
Qy 1201 SSGOSSKTEHWSSENSTPSSNAKRONLHPSSAQSRGQPKAATCKVSSINQETIQ 1260  
Db 1201 SSGOSSKTEHWSSENSTPSSNAKRONLHPSSAQSRGQPKAATCKVSSINQETIQ 1260  
Qy 1261 TYCVEDTPICFRCSLSUSSLSAEDICGNCQTQEADSAANTLQIAIEIKKIGTRSAEDPV 1320  
Db 1261 TYCVEDTPICFRCSLSUSSLSAEDICGNCQTQEADSAANTLQIAIEIKKIGTRSAEDPV 1320  
Qy 1321 SEVPAVSOHPTKSSRLQGSLSASEARHKA:VFSSGAKSPSKGACTPKSPPEHYVQET 1380  
Db 1321 SEVPAVSOHPTKSSRLQGSLSASEARHKA:VFSSGAKSPSKGACTPKSPPEHYVQET 1380  
Qy 1381 PLMFSTRCTSVSSLDSPFSRISASVQSEPCSGMVSGI:ISPSDLDPSPGQTMPPSRSKTPP 1440  
Db 1381 PLMFSTRCTSVSSLDSPFSRISASVQSEPCSGMVSGI:ISPSDLDPSPGQTMPPSRSKTPP 1440  
Qy 1441 PPQTAQTKREVKNKAPTAKESGPKQAANVAAVQVVOVLDPADTLHFATESPDGF 1500  
Db 1441 PPQTAQTKREVKNKAPTAKESGPKQAANVAAVQVVOVLDPADTLHFATESPDGF 1500  
Qy 1501 SCSSLSALSILDEPFIQKVELR:IMPPVQENDNGNETESQPKESNENQKAEKTI:DS 1560  
Db 1501 SCSSLSALSILDEPFIQKVELR:IMPPVQENDNGNETESQPKESNENQKAEKTI:DS 1560  
Qy 1561 KDLLDDSDDDDI:ELIEBCIIISAMPTKSKRAKPAQATASKLPPPVARKPSQLPVYKLLPS 1620  
Db 1561 KDLLDDSDDDDI:ELIEBCIIISAMPTKSKRAKPAQATASKLPPPVARKPSQLPVYKLLPS 1620  
Qy 1621 QNRLOPKHVSFTPGDDMPRVYCVGTP:INFSTATSLDITBSPNELAAGBVGREGQA 1680  
Db 1621 QNRLOPKHVSFTPGDDMPRVYCVGTP:INFSTATSLDITBSPNELAAGBVGREGQA 1680  
Qy 1681 SGFEKEDTITPTEGRSTDEAGGKTSSTV:PELDDNKAEGDILAEICINSAMPKGSHPK 1740  
Db 1681 SGFEKEDTITPTEGRSTDEAGGKTSSTV:PELDDNKAEGDILAEICINSAMPKGSHPK 1740  
Qy 1741 FRVKIMDVOQASASSANQKLDGKKKPTSPVKPI:PONTEYTRVRKNADSNNLN 1800  
Db 1741 FRVKIMDVOQASASSANQKLDGKKKPTSPVKPI:PONTEYTRVRKNADSNNLN 1800  
Qy 1801 AERFVSDNKD:SKKQNLKNNKDFNDKLPNNEDVRGSPFADSPHHTPIEGTPYCFSRND 1860  
Db 1801 AERFVSDNKD:SKKQNLKNNKDFNDKLPNNEDVRGSPFADSPHHTPIEGTPYCFSRND 1860  
Qy 1861 SLSLDDDDDDVLSRKAELRKA:KENKESEAKVTSHTLTSNOQANKTQATKOPINR 1920  
Db 1861 SLSLDDDDDDVLSRKAELRKA:KENKESEAKVTSHTLTSNOQANKTQATKOPINR 1920  
Qy 1921 GQPKPILQKSTFPQSSKIDPDRGAATDEKLQNFALIENTPVCFSHNSLSLSDIDQENN 1980  
Db 1921 GQPKPILQKSTFPQSSKIDPDRGAATDEKLQNFALIENTPVCFSHNSLSLSDIDQENN 1980  
Qy 1981 NKENPIKETEPDPSQGEPSKQASGYAPKSFHVETDTPVCFSRNSSLSSLSIDSEDDLLQ 2040

Db 1981 NKENPIKETEPDPSQGEPSKQASGYAPKSFHVETDTPVCFSRNSSLSSLSIDSEDDLLQ 2040  
Qy 2041 ECISAMPKKKKPRLKGDNEKHSFRNMGGLGSDLTLDLKDIO:RDPSEHGLSPDSNFD 2100  
Db 2041 ECISAMPKKKKPRLKGDNEKHSFRNMGGLGSDLTLDLKDIO:RDPSEHGLSPDSNFD 2100  
Qy 2101 WKATQEGANS:TVSSILHQAAAAACLSROASDSDSLILSKSGISLGSPPHLTPDDEKXPFT 2160  
Db 2101 WKATQEGANS:TVSSILHQAAAAACLSROASDSDSLILSKSGISLGSPPHLTPDDEKXPFT 2160  
Qy 2161 SNKQPRILKPEKSTLETKKIESESKGKGGKYYKSLITGKVRNSBEI:SGOMKQPOAN 2220  
Db 2161 SNKQPRILKPEKSTLETKKIESESKGKGGKYYKSLITGKVRNSBEI:SGOMKQPOAN 2220  
Qy 2221 MPSISRGCTMTHI:PGVRNSSSSTSPVSKGPPPLKTPASKSPSEGTATTSPRGAKPVS 2280  
Db 2221 MPSISRGCTMTHI:PGVRNSSSSTSPVSKGPPPLKTPASKSPSEGTATTSPRGAKPVS 2280  
Qy 2281 ELSPVARQTSQIGGSKAPSRSGSRDSTPSRPAQOPLSRPIQSPGRNS:ISPGRNGISPPN 2340  
Db 2281 ELSPVARQTSQIGGSKAPSRSGSRDSTPSRPAQOPLSRPIQSPGRNS:ISPGRNGISPPN 2340  
Qy 2341 KLSQLPRTSSPSTASTKSSGSKMYSYTPGROMSQOQLTKOTGLSKNASSIPRSESASKG 2400  
Db 2341 KISQLPRTSSPSTASTKSSGSKMYSYTPGROMSQOQLTKOTGLSKNASSIPRSESASKG 2400  
Qy 2401 LNQMNGNANKVELSRMSTKSSGSEDRSERPVLVRQSTFIKEAPSP:TLRKLRESA 2460  
Db 2401 LNQMNGNANKVELSRMSTKSSGSEDRSERPVLVRQSTFIKEAPSP:TLRKLRESA 2460  
Qy 2461 SFESLSPSSRRPASPTRSOAQTPLVSLPMSLTHSSVQAGGWRKLP:PKMLSPITIEYNDG 2520  
Db 2461 SFESLSPSSRRPASPTRSOAQTPLVSLPMSLTHSSVQAGGWRKLP:PKMLSPITIEYNDG 2520  
Qy 2521 RPAKHDIARSHSSPRLPINRSGTWKRHSKHSSILPRVSTWRTGSSSSILSASSES 2580  
Db 2521 RPAKHDIARSHSSPRLPINRSGTWKRHSKHSSILPRVSTWRTGSSSSILSASSES 2580  
Qy 2581 SEKAKSEKHVNSISGTSKQKENVSAKQWTRKIKENEFSPNTNSTQTVSSGATNGAES 2640  
Db 2581 SEKAKSEKHVNSISGTSKQKENVSAKQWTRKIKENEFSPNTNSTQTVSSGATNGAES 2640  
Qy 2641 KTLIYQWAPAVSKTEDVWVRIEDCPINNPRSGRSPGTGNTPPVIDSVSEKANENIKDSKN 2700  
Db 2641 KTLIYQWAPAVSKTEDVWVRIEDCPINNPRSGRSPGTGNTPPVIDSVSEKANENIKDSKN 2700  
Qy 2701 QAKQNVNGSVPMRTVGLNRLNSFIOVDAPDQXGTEIKPGQNNPVPVSETNESSI:VERT 2760  
Db 2701 QAKQNVNGSVPMRTVGLNRLNSFIOVDAPDQXGTEIKPGQNNPVPVSETNESSI:VERT 2760  
Qy 2761 PFSSSSSKHSSPBGTVAAARVTPFNYPNPSPRKSADSTSA:RPSQIPTPVANNNTKRSK 2820  
Db 2761 PFSSSSSKHSSPBGTVAAARVTPFNYPNPSPRKSADSTSA:RPSQIPTPVANNNTKRSK 2820  
Qy 2821 DSTSSGTSQSPKRHS:GSGYLVTSV 2843  
Db 2821 DSTSSGTSQSPKRHS:GSGYLVTSV 2843

## RESULT 7

US-10-267-502-370  
; Sequence 370, Application US/10267502  
; Publication No. US20040071700A1  
; GENERAL INFORMATION:  
; APPLICANT: Kim, Jaeseob  
; APPLICANT: Galant, Ron  
; TITLE OF INVENTION: Obesity Linked Genes  
; FILE REFERENCE: LSD-07416  
; CURRENT APPLICATION NUMBER: US/10/267,502  
; CURRENT FILING DATE: 2003-01-27  
; NUMBER OF SEQ ID NOS: 439  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 370

Query Match 99.4%; Score 14484.5; DB 12; Length 2844;  
Best Local Similarity 99.5%; Pred No. 0;  
Matches 2830; Conservative 3; Mismatches 10; Indels 1; Gaps 1;  
US-10-267-502-370

1 MAAASYDQLLKQVEALKWENSLRQLEEDNSNHLTKLETEASNKKEVLKQLGSIEDAM 60  
1 MAAASYDQLLKQVEALKWENSLRQLEEDNSNHLTKLETEASNKKEVLKQLGSIEDAM 60  
61 ASSGOIDLLERKELELNDSSNPPGVKLRSKMSLRSGREGSVSRSGECPVPMGSPPR 120  
61 ASSGOIDLLERKELELNDSSNPPGVKLRSKMSLRSGREGSVSRSGECPVPMGSPPR 120  
121 RGFVNGSRSTGYLBELEKERSLILLADLDKBEKDWYQAQQLTKRIDSPLTFENFSL 180  
121 RGFVNGSRSTGYLBELEKERSLILLADLDKBEKDWYQAQQLTKRIDSPLTFENFSL 180  
181 QDMLTRRQLEYEARQIRVAMEEQLGTCQDMEXRAQRRIARQQIEKDILRTRQLLOSAT 240  
181 QDMLTRRQLEYEARQIRVAMEEQLGTCQDMEXRAQRRIARQQIEKDILRTRQLLOSAT 240  
241 EAERSSONKHETGSHDAERONEGQVGEINMATSGNGOGSTTRMDHETASVLSGSSTHSA 300  
241 EAERSSONKHETGSHDAERONEGQVGEINMATSGNGOGSTTRMDHETASVLSGSSTHSA 300  
301 PRLTSHLGTKEVMYVLSLMSLGTDXDMDSTLLAMSSQDSCISMRQSGCLPILLIQLL 360  
301 PRLTSHLGTKEVMYVLSLMSLGTDXDMDSTLLAMSSQDSCISMRQSGCLPILLIQLL 360  
361 HGNKDVSLLNGSRGSKAARASAAALHNIHISQPDCKGREGREIRVLHLQIRAYCETC 420  
361 HGNKDVSLLNGSRGSKAARASAAALHNIHISQPDCKGREGREIRVLHLQIRAYCETC 420  
421 WEMQBAHEFGMDQKNPMPAPVEHOICPAVCVLMKLSFDEEHRHAMELGGLOAIAELLQ 480  
421 WEMQBAHEFGMDQKNPMPAPVEHOICPAVCVLMKLSFDEEHRHAMELGGLOAIAELLQ 480  
481 VDCMYGLTNDHYSITLRYAGMALTNLTFFGVANKATLCSMGKGCVRALVAQLKSESD 540  
481 VDCMYGLTNDHYSITLRYAGMALTNLTFFGVANKATLCSMGKGCVRALVAQLKSESD 540  
541 QQVIASVLNLSWRADVNSKKTILREVGSVKALMECALEVKKESTLKSVLNLSAHC 600  
541 QQVIASVLNLSWRADVNSKKTILREVGSVKALMECALEVKKESTLKSVLNLSAHC 600  
601 ENKADICAVDGAFLVGLTVRSQNTLAIIESGGGILRNVSLSIATNEDHRQILRENN 660  
601 ENKADICAVDGAFLVGLTVRSQNTLAIIESGGGILRNVSLSIATNEDHRQILRENN 660  
661 CLQTLLOHLKSHSLTIVSNACGTLNLSARNPKDQALWDGAVSMKLNLIHSHKMTIAM 720  
661 CLQTLLOHLKSHSLTIVSNACGTLNLSARNPKDQALWDGAVSMKLNLIHSHKMTIAM 720  
721 GSAALRNLMANRPARYKDANIMSPGSSLPFLVRKQKALEAELDAQHLSETFDNLS 780  
721 GSAALRNLMANRPARYKDANIMSPGSSLPFLVRKQKALEAELDAQHLSETFDNLS 780  
781 PKASHRSKQHKOSLYGDVYFDRNRHDDNRSNFTNGNTVLSPLYNTTVLPSSSSRGS 840  
781 PKASHRSKQHKOSLYGDVYFDRNRHDDNRSNFTNGNTVLSPLYNTTVLPSSSSRGS 840  
841 LDSRSSEKORSLEBERGILGNVHPATENPGTSSKRGLOISITAAQIAKMEEVSAIHTS 900  
841 LDSRSSEKORSLEBERGILGNVHPATENPGTSSKRGLOISITAAQIAKMEEVSAIHTS 900  
901 QEDRSSGTTTELHCVTDERNALRRSAAHTHSNTYNTFTKSENSTCSNMPYAKLEYKRSS 960  
901 QEDRSSGTTTELHCVTDERNALRRSAAHTHSNTYNTFTKSENSTCSNMPYAKLEYKRSS 960

Qy 961 NDSLNSVSSSDGYGKRGQMKPSIESYSEDDSKFCSGYQYPADLAHKHISANHMDNDGE 1020  
Db 961 NDSLNSVSSNDGYGKRGQMKPSIESYSEDDSKFCSGYQYPADLAHKHISANHMDNDGE 1020  
Qy 1021 LDTPIYSLKYSDQNLNQRSPSONERWARPKHIIIEDEIKQSEQRQSRNQSTTYPVYTE 1080  
Db 1021 LDTPIYSLKYSDQNLNQRSPSONERWARPKHIIIEDEIKQSEQRQSRNQSTTYPVYTE 1080  
Qy 1081 STDKHLKFPQHGQOECVSPYRSRGANGSETNRVGSNHCINQNVQSGLCOEDDYEDDKP 1140  
Db 1081 STDKHLKFPQHGQOECVSPYRSRGANGSETNRVGSNHCINQNVQSGLCOEDDYEDDKP 1140  
Qy 1141 TNSYERYSEBEEHEERPTNYSIKYNEEKRVDPQIDYSLKYATDIPSSQKQSFSPFSKS 1200  
Db 1141 TNSYERYSEBEEHEERPTNYSIKYNEEKRVDPQIDYSLKYATDIPSSQKQSFSPFSKS 1200  
Qy 1201 SSGOSSKTEHMSSENSTSTPSSNAKRONQLHPSSAQSRSGQPKAATCKVSSINQETIQ 1260  
Db 1201 SSGOSSKTEHMSSENSTSTPSSNAKRONQLHPSSAQSRSGQPKAATCKVSSINQETIQ 1260  
Qy 1261 TYCVEDTPICFSCSSISLSSAEDIEGQNTTQEADSAANTLOIAELKEKIGTRSAEDPV 1320  
Db 1261 TYCVEDTPICFSCSSISLSSAEDIEGQNTTQEADSAANTLOIAELKEKIGTRSAEDPV 1320  
Qy 1321 SEVP-AVSOHPRTKSSRLQSSLSSESARHAKAVEFSPSGAKSPSKGAQTSPPEHYVQE 1379  
Db 1321 SEVPSSVHSTLETKSSRLQSSLSSESARHAKAVEFSPSGAKSPSKGAQTSPPEHYVQE 1380  
Qy 1380 TPLMFSCTSVSSLDSPESRSIASSVQSEPCSGWVGIIISPSDLPSPGQTMPPSRKTP 1439  
Db 1381 TPLMFSCTSVSSLDSPESRSIASSVQSEPCSGWVGIIISPSDLPSPGQTMPPSRKTP 1440  
Qy 1440 PPPQTAQTREVPKNAPTAEKRESGPKQAAVNAAVQRVQVLPDADTLHFAETESTPDG 1499  
Db 1441 PPPQTAQTREVPKNAPTAEKRESGPKQAAVNAAVQRVQVLPDADTLHFAETESTPDG 1500  
Qy 1500 FSCSSISLSSALSLDEPFIQKVELRIMPVOENNGNTESEBQPKESNENQKAEKTIIDS 1559  
Db 1501 FSCSSISLSSALSLDEPFIQKVELRIMPVOENNGNTESEBQPKESNENQKAEKTIIDS 1560  
Qy 1560 EKDLDDSDDDDEILBECIIISAMPTKSRKAKKPAOTASKLPPPVARKESQLPVYKLLP 1619  
Db 1561 EKDLDDSDDDDEILBECIIISAMPTKSRKAKKPAOTASKLPPPVARKESQLPVYKLLP 1620  
Qy 1620 SQNRLOPKHVSFTPGDDMPRVYVEGTPIINFSTATSLDITIESPPNELAAGEVVGGA 1679  
Db 1621 SQNRLOPKHVSFTPGDDMPRVYVEGTPIINFSTATSLDITIESPPNELAAGEVVGGA 1680  
Qy 1680 QSGFEKRDITPTBGRSTDEAQGGKTSVTIPELDDNKABEGDILAECSINAMPKGSKH 1739  
Db 1681 QSGFEKRDITPTBGRSTDEAQGGKTSVTIPELDDNKABEGDILAECSINAMPKGSKH 1740  
Qy 1740 PFRVYKIMDQVQASASSAPNKNQLDGKKKXPTSPVKPIPQNTYRTRVRKNADSKNL 1799  
Db 1741 PFRVYKIMDQVQASASSAPNKNQLDGKKKXPTSPVKPIPQNTYRTRVRKNADSKNL 1800  
Qy 1800 NAERVFSNCKSKQNLKNNSKDFNDKLPNNEDRVGSPAFDSPHHYTPTEGTPYCFSRN 1859  
Db 1801 NAERVFSNCKSKQNLKNNSKDFNDKLPNNEDRVGSPAFDSPHHYTPTEGTPYCFSRN 1860  
Qy 1860 DLSLSLDFDDDDVLSREKAELEKAKENKESAKVTSHTELTNQQSANKTQAIKQPIN 1919  
Db 1861 DLSLSLDFDDDDVLSREKAELEKAKENKESAKVTSHTELTNQQSANKTQAIKQPIN 1920  
Qy 1920 RGQPKPILOKQSTPPQSSKOIPORGAATDEKLNQFAIENTPVCFSHNSLSLSLSDIDQEN 1979  
Db 1921 RGQPKPILOKQSTPPQSSKOIPORGAATDEKLNQFAIENTPVCFSHNSLSLSLSDIDQEN 1980  
Qy 1980 NKNENEPKETEPPDSOGEPKQASGVAPKSFHVEDTPTVCFSRNSLSLSLSDIDSEDDL 2039  
Db 1991 NKNENEPKETEPPDSOGEPKQASGVAPKSFHVEDTPTVCFSRNSLSLSLSDIDSEDDL 2040  
Qy 2040 QECISSAMPKKKPKSRUKGDNKXHSRPMGGIIGEDLTLDKDIQRPDSEHGLSPDSNF 2099

```

Db 2041 QECISSAMPKKKYSRLKGDNEKSPRNMGILGEDJTLDKIQRPDSEHGSLSPDSENF 2100
Qy 2100 DWKAIQEGANSIVSLHQAQAAAACLSQAASDSDSIILSLKSGISLGSPFHLTPDQBEKPF 2159
Db 2101 DWKAIQEGANSIVSLHQAQAAAACLSQAASDSDSIILSLKSGISLGSPFHLTPDQBEKPF 2160
Qy 2160 TSNKGPRILKPEKSTLETKIESESGIKGKVKYKSLITGKVRNSEISGQMKPQLOA 2219
Db 2161 TSNKGPRILKPEKSTLETKIESESGIKGKVKYKSLITGKVRNSEISGQMKPQLOA 2220
Qy 2220 NMPISIRGRTHIFGVNRSNSSSTSPVSKKGPPLKTPASKSPSEGQTATTSRGAKEPSVK 2279
Db 2221 NMPISIRGRTHIFGVNRSNSSSTSPVSKKGPPLKTPASKSPSEGQTATTSRGAKEPSVK 2280
Qy 2280 SELSPVARQTSQIGSSKAPRSRSDSTPSRPAQQLSRPIQSPGRNSISPORNGISPP 2339
Db 2281 SELSPVARQTSQIGSSKAPRSRSDSTPSRPAQQLSRPIQSPGRNSISPORNGISPP 2340
Qy 2340 NKLQOLPRTSPSTASTKSSGSGRMSYTSRQMSQQLTKQTGLSKVASSISPRSESASK 2399
Db 2341 NKLQOLPRTSPSTASTKSSGSGRMSYTSRQMSQQLTKQTGLSKVASSISPRSESASK 2400
Qy 2400 GLNOMNNGANGKVKELSRMSSTKSSGSGRMSYTSRQMSQQLTKQTGLSKVASSISPRSESASK 2459
Db 2401 GLNOMNNGANGKVKELSRMSSTKSSGSGRMSYTSRQMSQQLTKQTGLSKVASSISPRSESASK 2460
Qy 2460 ASPESLSPSSRPASPTSOQTPVLSPLDMSLSTHSSVQAGGWKLPNLSPTTIEYND 2519
Db 2461 ASPESLSPSSRPASPTSOQTPVLSPLDMSLSTHSSVQAGGWKLPNLSPTTIEYND 2520
Qy 2520 GRPAKRHDIAHSHSESPRLPINRSGTWKREHSHSSSLPRVSTWRTGSSSSILSASSE 2579
Db 2521 GRPAKRHDIAHSHSESPRLPINRSGTWKREHSHSSSLPRVSTWRTGSSSSILSASSE 2580
Qy 2580 SSEKAKSEDEKHVNSISGTSQSKENOVSAKGTWRKIKENEFSTNSTQTSVSSGATNGAE 2639
Db 2581 SSEKAKSEDEKHVNSISGTSQSKENOVSAKGTWRKIKENEFSTNSTQTSVSSGATNGAE 2640
Qy 2640 SKTLIYQMAPAVSKTEDVWVRIEDCPINNPRSGRSPGTGNTPPVIDSVSEKANPNIKDSKD 2699
Db 2641 SKTLIYQMAPAVSKTEDVWVRIEDCPINNPRSGRSPGTGNTPPVIDSVSEKANPNIKDSKD 2700
Qy 2700 NOAKQNVGSGVPWRTVGLNRLNSFTQVADPKGTEIKPGQNNPVPVSETNESSIVER 2759
Db 2701 NOAKQNVGSGVPWRTVGLNRLNSFTQVADPKGTEIKPGQNNPVPVSETNESSIVER 2760
Qy 2760 TPFSSSSSKHSSPSGTVAAARVTPFNVPNPRKSSADSTARSQIPTPVNNNTKKRDSK 2819
Db 2761 TPFSSSSSKHSSPSGTVAAARVTPFNVPNPRKSSADSTARSQIPTPVNNNTKKRDSK 2820
Qy 2820 TDSTESSGTSQPKRHSGSYLVTSV 2843
Db 2821 TDSTESSGTSQPKRHSGSYLVTSV 2844

```

RESULT 8

```

US-10-267-502-372
; Sequence 372, Application US/10267502
; Publication No. US2004007100A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jaeseob
; APPLICANT: Gallant, Ron
; TITLE OF INVENTION: Obesity Linked Genes
; FILE REFERENCE: LSD-07416
; CURRENT APPLICATION NUMBER: US/10/267, 502
; CURRENT FILING DATE: 2003-01-27
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 372
; LENGTH: 2845
; TYPE: PR

```

US-10-267-502-372

```

Query Match 90.5%; Score 13178; DB 12; Length 2845;
Best Local Similarity 90.2%; Pred. No. 0;
Matches 2572; Conservative 111; Mismatches 155; Indels 12; Gaps 10;

Qy 1 MAAASVDQLLKQVEALKMENSNIHQLEEDNSNHLTKLETEASNMEKVLKQIGSTEDSAM 60
Db 1 MAAASVDQLLKQVEALKMENSNIHQLEEDNSNHLTKLETEASNMEKVLKQIGSTEDSAM 60

Qy 61 AASGQIDLLERLKEINLDSSNPPFGVKLSKMSLSRYSYSGRSGSVSRSGECSPVPMGSPFR 120
Db 61 -TSGQIDLLERLKEINLD-SNPPFGVKLSKMSLSRYSYSGRSGSVSRSGECSPVPMGSPFR 118

Qy 121 RGFVNGSRESTGYLBELEKERSLLIADLDKBEKEKDWYYAQIQLTKRIDSPLTENFSL 180
Db 119 RTFVNGSRESTGYLBELEKERSLLIADLDKBEKEKDWYYAQIQLTKRIDSPLTENFSL 178

Qy 181 QTDMTRRQLEYEARQIRVAMEEQLGTCODMEKRAORRIARIQOIEKDTLRIRQLQSOAT 240
Db 179 QTDMTRRQLEYEARQIRVAMEEQLGTCODMEKRAORRIARIQOIEKDTLRIRQLQSOAA 238

Qy 241 EAERSSQNHETGSHDAERQNEGOVGINWATSGNGQSTTRMDHETASVLSSSSTHSA 300
Db 239 EAERSSQNRHDAASHEAGRCHEGHAESNTAASSSQSPATRVVDHETASVLSSSSTHSA 298

Qy 301 PRLTSHLGTKEVMYISLLSMLGTHDKDDMSRTLLAYSSSDSCISMEQSCPLLIQLL 360
Db 299 PRLTSHLGTKEVMYISLLSMLGTHDKDDMSRTLLAYSSSDSCISMEQSCPLLIQLL 358

Qy 361 HGNDKDSVLLGNRSGSKARARASAAHLNIIHSPQDDKGRREIRVLHLEQIRAYCETC 420
Db 359 HGNDKDSVLLGNRSGSKARARASAAHLNIIHSPQDDKGRREIRVLHLEQIRAYCETC 418

Qy 421 NEWQEAHPGDGDDKQNPAPVEHQICPAVCVLNKLSPDEHRHAMELGLQIAELLQ 480
Db 419 NEWQEAHPGDGDDKQNPAPVEHQICPAVCVLNKLSPDEHRHAMELGLQIAELLQ 478

Qy 481 VDCEMYGLTNDHYSITLRRYAGMALTNLTFGDVANKATLCMKCMKALVAQLKSEEDL 540
Db 479 VDCEMYGLTNDHYSITLRRYAGMALTNLTFGDVANKATLCMKCMKALVAQLKSEEDL 538

Qy 541 QOVIASVLRNLSRADVNSKKTIREVGSVKALMECALVEKKESTLKSVALNWLNSAHT 600
Db 539 QOVIASVLRNLSRADVNSKKTIREVGSVKALMECALVEKKESTLKSVALNWLNSAHT 598

Qy 601 ENKADICAVDGCALAFVLGTLTYRSOTNTLAIIESGGGILRNVSLLIATNEDHROIAREN 660
Db 599 ENKADICAVDGCALAFVLGTLTYRSOTNTLAIIESGGGILRNVSLLIATNEDHROIAREN 658

Qy 661 CLQTLLOHLKSHSLITIVSNACGTLNLSARNPKDOEALWDMGAVSMLKNLIHSHKMIAM 720
Db 659 CLQTLLOHLKSHSLITIVSNACGTLNLSARNPKDOEALWDMGAVSMLKNLIHSHKMIAM 718

Qy 721 GSAALRNLMANRPAPKYKDNIMSPGSSLSPLSHVRKOKALEAELEDAOHLSETFNINLS 780
Db 719 GSAALRNLMANRPAPKYKDNIMSPGSSLSPLSHVRKOKALEAELEDAOHLSETFNINLS 778

Qy 781 PKASHRSKORHKQSLGYDYVFDITNRHDDNRSDNFNTGMTVLSPLYLNTTVLPSSSSSRGS 840
Db 779 PKASHRSKORHKQSLGYDYVFDITNRHDDNRSDNFNTGMTVLSPLYLNTTVLPSSSSSRGS 838

Qy 841 LDSSRSKORSERERGLGNYHPATENGTSSKRGLOISTTAAQIAKYMEEVSAIHTS 900
Db 839 LDSSRSKORSERERGLGNYHPATENGTSSKRGLOISTTAAQIAKYMEEVSAIHTS 898

Qy 901 QEDRSSGSTTTEHCVTDERNALRSSAAHSTHNTYNTFKSNGSNRTCSMPYAKLEYKRSS 960
Db 899 QEDRSSGSTTTEHCVTDERNALRSSAAHSTHNTYNTFKSNGSNRTCSMPYAKLEYKRSS 958

Qy 961 NDSLNSVSSSDGYCKRQMKPSIESYSEDDSEKFCYGOYPADLAHSHANEMDDNDGE 1020
Db 959 NDSLNSVSSSDGYCKRQMKPSIESYSEDDSEKFCYGOYPADLAHSHANEMDDNDGE 1020

```

2098	Qy	NPDWAIQEGANSIYVSSLHO--AAAACILSQRASSSDSDTILSLKSGILSGSPHLTPDQBE	2156
2097	Db	NPDWAIQEGANSIYVSSLHQAAAAACILSQRASSSDSDTILSLKSGILSGSPHLTPDQBE	2156
2157	Qy	KPFTSNKGPRLKPEKSTLETKKIESBSKGIKGGKVYKSLITGKVRNSISISGOMKOP	2216
2157	Db	KPFTSNKGPRLKPEKSTLEAKKIESENKGIKGGKVYKSLITGKIRNSBISSOMKOP	2216
2217	Qy	LOANNPSTSRGHTMTIHIPVRNNSSSSTSPVSKKGPFLLKTPASKSPSEGOATITSPRGAKP	2276
2217	Db	LPTNMPSTSRGHTMTIHIPGRNNSSSSTSPVSKKGPFLLKTPASKSPSEGGCATITSPRGTKP	2276
2277	Qy	SVKSELSPVAROTSGTQSSKAPASGSRSDTSPSPAQOPLSRPIOSPGRNSISPGRNGI	2336
2277	Db	ACKSELSPIRQTSQISGNSKSSRSGSRSDTSPSRPTQPLSRPMQSPGRNSISPGRNGI	2336
2337	Qy	SPPNKLSQLPRTSSPSTASTKSSGGGKMSYITSPGRQMSQOONLTQTCLSKNASSIIPRSES	2396
2337	Db	SPPNKLSQLPRTSSPSTASTKSSGGGKMSYITSPGRQLSQOONLTQASLSKNASSIIPRSES	2396
2397	Qy	ASKGLNQMNNGANKKVELSEMSSTKSSGSESDSERPVLVRQSTFIKEAPSPTLRRKL	2456
2397	Db	ASKGLNQMNNGNSKKVELSEMSSTKSSGSESDSERPALVRQSTFIKEAPSPTLRRKL	2456
2457	Qy	ESASFPESLPSRRPASPTRSOAQTPVLSPSIPDMSLSTHSSVQACGWKRLPNLSPPTIE	2516
2457	Db	ESASFPESLPSRRPDSPTRSOAQTPVLSPSIPDMSLSTHPSVQACGWKRLPNLSPPTIE	2516
2517	Qy	YNDGRPAKHDTLARSHESPSRLPINRSGTWRKHSKHSSSLPRVSTWRTTSSSSSILSA	2576
2517	Db	YNDGRPTKHDILARSHESPSRLPINRACTWRKHSKHSSSLPRVSTWRTGSSSSILSA	2576
2577	Qy	SSESSEKAKSEDEKHVNSISGTQKQSKENOVASKGTWRKIKEHNEFSTNSTQTSVVSGATN	2636
2577	Db	SSESSEKAKSEDERHVSSPAPQMKENQVPTKGTWRKIKESDISPTGMAQSASGAAS	2636
2637	Qy	GAESKTLIIQMAPAVSKTDEWVRVEDCPINNPRSGRSPGTGTPPVIDSVSEKANPNIXD	2696
2637	Db	GAESKPLIIQMAPPVSKTEDWVRVEDCPINNPRSGRSPGTGTPPVIDSVSEKSSIXD	2696
2697	Qy	SKDQ--AKQNVGNSVPMTVUGLENRLNPSIFQVDAPQOKGTEIIPQGNPNVPVSETNE	2753
2697	Db	SKDSKDTHGQKSYGSGS--PVQTVGLETRLNSFVQVEAPQKGTEAKPGGSPNPSVIAETA	2755
2754	Qy	SSIVERTPTSSSSSSKHSSPBGTVAAARVTPFNYNPSPKSSADSTGARPQIETPPVNNNT	2813
2756	Db	TCIAERTPTSSSSSSKHSSPBGCTVAARVTPFNYNPSPKSSADSTGARPQIETPVSNT	2815
2814	Qy	KKRDSTDTSESSGTOQSPKRHSGSYLVTSV	2843
2816	Db	KKRDSTKITDESSGAQSPKRHSGSYLVTSV	2845

US-09-987-482-2

US-09-367-402-2  
: Sequence 2, Application US/09987482

Publication No. US20020184656A1

GENERAL INFORMATION:

APPLICANT: BHANDARI, POONAM

APPLICANT: SHASHIDHARA, L.S.

; TITLE OF INVENTION: IN VIVO ASSAY SYSTEM FOR SC

**TITLE OF INVENTION: DRUGS AND OTHER SUBSTANCES**

; FILE REFERENCE: 056859-0134

CURRENT APPLICATION NUMBER: US/09/987,482

; CURRENT FILING DATE: 2002-03-21  
 ; STRIPED OF CEO ID NOS: 3

```

; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patcnt In War

```

; SOFTWARE: F  
; CEO ID NO: 3

```

; SEQ ID NO 2
; LENGTH: 912

```

```

; LENGTH: 512
; TYPE: PRT

```

ORGANISM: Homo sapiens

US-09-987-482-2

```

Query Match      32.5%; Score 4738; DB 9; Length 912;
Best Local Similarity 100.0%; Pred. No. 3.3e-259;
Matches 912; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 959 SSNDLSNVSSDGYGKRGKMPKSISSYEDDSKFCSCYQYPADLAHKIHSANHMDND 1018
Db 1 SSNDLSNVSSDGYGKRGKMPKSISSYEDDSKFCSCYQYPADLAHKIHSANHMDND 60

Qy 1019 GELDTPIYSLKYSDEQLNSGRSPSONERWAPKHIIEDEIKQSQQRNQSTTYPY 1078
Db 61 GELDTPIYSLKYSDEQLNSGRSPSONERWAPKHIIEDEIKQSQQRNQSTTYPY 120

Qy 1079 TESTDDXELAFQHPFGQECVSPYRSGANGSTNRVGNHGINQNVOSLCOEDDYEDD 1138
Db 121 TESTDDXELAFQHPFGQECVSPYRSGANGSTNRVGNHGINQNVOSLCOEDDYEDD 180

Qy 1139 KPTNYISRYSEEBQHEEERPTNYISKYBEXKHVDQPIDYSLKYATDIPSSQXQSFSPS 1198
Db 181 KPTNYISRYSEEBQHEEERPTNYISKYBEXKHVDQPIDYSLKYATDIPSSQXQSFSPS 240

Qy 1199 KSSSGQSSKTEHMSSESSENTSTPSSNAKRONLHPSSAQSRGPOKAAATCKVSSINQET 1258
Db 241 KSSSGQSSKTEHMSSESSENTSTPSSNAKRONLHPSSAQSRGPOKAAATCKVSSINQET 300

Qy 1259 IQYCVYEDTPICFSRCSLSLSAEDIEICNQTTOEADSANTLQIAEIKKIGTRSAED 1318
Db 301 IQYCVYEDTPICFSRCSLSLSAEDIEICNQTTOEADSANTLQIAEIKKIGTRSAED 360

Qy 1319 PVSEVPVAVSHPRTKSRLOGLSSLSSESABKHAVERSSGNAKSPKSGAOTPKSPPHYVQ 1378
Db 361 PVSEVPVAVSHPRTKSRLOGLSSLSSESABKHAVERSSGNAKSPKSGAOTPKSPPHYVQ 420

Qy 1379 ETPLMFSCRTSVSSLDSEFERSIASSVQSEPCGMVSGIISPSDLDPSPQCTMPSPRSKT 1438
Db 421 ETPLMFSCRTSVSSLDSEFERSIASSVQSEPCGMVSGIISPSDLDPSPQCTMPSPRSKT 480

Qy 1439 PPPPQTAQTKREVKNKAPTAKRESGPKQAANAQVQVLPADTLLHPATSTPD 1498
Db 481 PPPPQTAQTKREVKNKAPTAKRESGPKQAANAQVQVLPADTLLHPATSTPD 540

Qy 1499 GFCSLSLSALSDEPPIQKDVLRIMPPVQENDNGNETESEOPKESNENQEKAEKTIID 1558
Db 541 GFCSLSLSALSDEPPIQKDVLRIMPPVQENDNGNETESEOPKESNENQEKAEKTIID 600

Qy 1559 SEXDLLDSDDDDIIELEECIIISAMPTKSRKAKPAQATASKLPPVARKPSQLPVYKLL 1618
Db 601 SEXDLLDSDDDDIIELEECIIISAMPTKSRKAKPAQATASKLPPVARKPSQLPVYKLL 660

Qy 1619 PSQNRLOPKHVGFPTGDDMPRYVCVEGTPINFSTATISDLTIESPPNELAAGEVRGG 1678
Db 661 PSQNRLOPKHVGFPTGDDMPRYVCVEGTPINFSTATISDLTIESPPNELAAGEVRGG 720

Qy 1679 AQSGEPEKRTDIPTEGRSTDEAQQGKTSVTIPELDDNKAEEGDILAEICINSAMPKGS 1738
Db 721 AQSGEPEKRTDIPTEGRSTDEAQQGKTSVTIPELDDNKAEEGDILAEICINSAMPKGS 780

Qy 1739 KPRVKIMQVQOQASASSAPKNOLDGKKKPTSPVXPIQNTYRTRVRKNADSKN 1798
Db 781 KPRVKIMQVQOQASASSAPKNOLDGKKKPTSPVXPIQNTYRTRVRKNADSKN 840

Qy 1799 LNAERFSDNKSQKQNLKNSKDFNDKLPNNEDVRGSAFDPSPHYTPIEGTPYCFGR 1858
Db 841 LNAERFSDNKSQKQNLKNSKDFNDKLPNNEDVRGSAFDPSPHYTPIEGTPYCFGR 900

Qy 1859 NDSLSSLDFFDD 1870
Db 901 NDSLSSLDFFDD 912

```

```

; GENERAL INFORMATION:
; APPLICANT: BHANDARI, POONAM
; APPLICANT: SHASHIDHARA, L.S.
; TITLE OF INVENTION: IN VIVO ASSAY SYSTEM FOR SCREENING AND VALIDATION OF
; TITLE OF INVENTION: DRUGS AND OTHER SUBSTANCES
; FILE REFERENCE: 056859-0134
; CURRENT APPLICATION NUMBER: US/09/987,482
; CURRENT FILING DATE: 2002-03-21
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 3
; LENGTH: 767
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-987-482-3

Query Match      26.7%; Score 3892; DB 9; Length 767;
Best Local Similarity 100.0%; Pred. No. 1.1e-211;
Matches 767; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAAASYDOLKQVAKWNSNLRQLEDNSNHLTKLETEASNKVELKQLQGSIRDEAM 60
Db 1 MAAASYDOLKQVAKWNSNLRQLEDNSNHLTKLETEASNKVELKQLQGSIRDEAM 60

Qy 61 ASSQIDLLERKELNLDSSNFPQVGLRSKMSLRSYSGREGSVSSRSGECPVPMGSFPR 120
Db 61 ASSQIDLLERKELNLDSSNFPQVGLRSKMSLRSYSGREGSVSSRSGECPVPMGSFPR 120

Qy 121 RGFVNGRESYGLLEELEKERSLLADLDKEEKKWYQAQLQNLTKRIDSLPTENFSL 180
Db 121 RGFVNGRESYGLLEELEKERSLLADLDKEEKKWYQAQLQNLTKRIDSLPTENFSL 180

Qy 181 QDTRRQLEYEAPQIRVAMEEQLTGQDMKRAQRARIQIIEKDILRIQLQSOAT 240
Db 181 QDTRRQLEYEAPQIRVAMEEQLTGQDMKRAQRARIQIIEKDILRIQLQSOAT 240

Qy 241 EAERSSQNKHETGSHDAERQNEGQVGEINWATSGNCGSTTRMDHETASVLSSTHSA 300
Db 241 EAERSSQNKHETGSHDAERQNEGQVGEINWATSGNCGSTTRMDHETASVLSSTHSA 300

Qy 301 PRRLTSLHGTVMVYSLLSMLGTHDKDMSRTLLANSSQDSCISMRQSGCLPLLIQLL 360
Db 301 PRRLTSLHGTVMVYSLLSMLGTHDKDMSRTLLANSSQDSCISMRQSGCLPLLIQLL 360

Qy 361 HGNDKDSVLLGNSRSGSEARASAAALHNTIHSOPDDKRGRRREIRVLHLEQIRAYCETC 420
Db 361 HGNDKDSVLLGNSRSGSEARASAAALHNTIHSOPDDKRGRRREIRVLHLEQIRAYCETC 420

Qy 421 WEMQEAHPGMDQKPMAPVVEHQICPVCVLMKLSFDEEHRHAMNELGLOAIAELLQ 480
Db 421 WEMQEAHPGMDQKPMAPVVEHQICPVCVLMKLSFDEEHRHAMNELGLOAIAELLQ 480

Qy 481 VDCMYGLTNDHYSITLRRYAGVALTNLTFTGDVANKATLCMKGCMRALVAQLKSEEDL 540
Db 481 VDCMYGLTNDHYSITLRRYAGVALTNLTFTGDVANKATLCMKGCMRALVAQLKSEEDL 540

Qy 541 QCVTASVLRNLSEADVNSKKTIREVGSVKALMECALEVKESTLKVLTGALNLSAHCT 600
Db 541 QCVTASVLRNLSEADVNSKKTIREVGSVKALMECALEVKESTLKVLTGALNLSAHCT 600

Qy 601 ENKADICAVDGCALAFVLTITYSQTNTLAIIESGGGILRNVSLSIATNEDHRQILRENN 660
Db 601 ENKADICAVDGCALAFVLTITYSQTNTLAIIESGGGILRNVSLSIATNEDHRQILRENN 660

Qy 661 CLQTLQHLKSHSLTIIVSNACGTLWNLSARNPKDQEAALMDMGAVMLKNIHSHKHMIAM 720
Db 661 CLQTLQHLKSHSLTIIVSNACGTLWNLSARNPKDQEAALMDMGAVMLKNIHSHKHMIAM 720

Qy 721 GSAALNLNANRPAPKYKDNIMSPGSSLPSPSLHVRKQKALEAELDAQ 767
Db 721 GSAALNLNANRPAPKYKDNIMSPGSSLPSPSLHVRKQKALEAELDAQ 767

```

[illegible]



```

QY 1810 DSKQNLKNNKOPNDKPNNEEDVRGSAFADSPHHYPTIEGTPYCFSRNDSLSLDFDD 1869
D 1464 DSSQOT-----RTRGDGALQSLCLTTPTEBAVYCF-----YDS 1496
QY 1870 DDVDLS-----REKABLRKAKENKSEAKVTSHTLTNSQOANKTOAIKOPIN---R 1920
D 1497 DEPPATAPPBRASAIPLRAKREKPAKR-----ETPSRAAQPATLPVR 1541
QY 1921 GQPKILQKQSTPPOSSKIDPDRGAATDEKLFQNTAENTVPCPSHNSLSLSDIDQENN 1980
D 1542 AQFLI-----VDTEPPCYSLTSSASSLSE----- 1566
QY 1981 NKNEPIKETEPDPSQGEPSKPOASGVAPKSFHVEDTPVCFPSRNSLSLSIDSEDDLIQ 2040
D 1567 -----PEA-----PQPANHARGPEQGSQD-----SSPSRAEBELIQ 1600
QY 2041 ECISSAMP-----KKKPSRLKGDNEKHSRPNMGGLGDLTLDLKDIOQPDSE-H 2090
D 1601 RCISLAMPRTTQVGGRRRKPRALRD-----IRPTEIT-----QKQEEVA 1643
QY 2091 GLSP--DSENFDMKAIQEGANSIVSSILHQAACLSRQASSDSLSILKSGISLGSPP 2148
D 1644 GSDPASLDSVWQAIQEGANSIVTLHQAAKASL--EASSESLSLSLVGSVAGSTL 1701
QY 2149 HLTPDQEKFTSNKGRILKPKGKSTLETYKI--ESESKIKGKVKYKSLITGKYR-SN 2206
D 1702 Q-----PSLKRKGKPAEAGGAWRPEKRTGTTSTK-----INGSRLFN 1740
QY 2207 SEISQWKQPLQANMPSISGRMTIHIPGVNRNSSTSPVSKGGLPKTPASKSPSEGT 2266
D 1741 GPEKAKTQXMMAGESTMLGRITVY-----SAGPASRTQSKG--ISGCTTTPKKTGS 1792
QY 2267 ATTPRGAKSVKSELSPVARQTSQICGSKAKRSRGRSDTSPRPAQPLRPISQGR 2326
D 1793 GTTQFETV-----TKAPSEQQRSSRLHRPKTSELALRHPPR 1831
QY 2327 NSISPRNGISPPNKLQSLPRTSP-----STASTKSG-----SGKMSVTSRQMSQQN 2377
D 1832 SATPARLAKTPSSSSQTSFASQPLRRRPLAIPGTGGLPGGSLVPKSPARAL---- 1887
QY 2378 LTKQTGLSKNASSI-----PRSESASKGLNQMNNGA--NKKVELSRMS 2420
D 1888 LAKQHTQKSPVRIFFMQRARRVPPPLARFPEPGRGAGAEPTPGARGSLGLVRMA 1947
QY 2421 STKSSGESDESRLVQSTFKAPSTLARKLEESAFESLSPSRPASPTBSQAQ 2480
D 1948 SARSSGES--SDRSGFRQLTFIKESFG--LLRRRSELSDASTASTQAAAPRGRPA 2004
QY 2481 TPVL--SPSLPDMSLSTHSSVQAGGRKLPPLNLSPTIYNDGRPAKRDHIAHSHSESPS 2537
D 2005 LPAVFLCSSRCDLURVSPQPLAQRSPQAKPLAPL-----APR-----RTSSESPS 2052
QY 2538 RLPINRSGTWKREKSHSSLSPLVSTWRTGSSSSLSASSESSEKAKSEDEKXVNSIS- 2596
D 2053 RLPIV-RASQRPETVKRYASLPHISVSRSDSAVSPTTQANATRRGSGDEARPLPRVAP 2111
QY 2597 -GTKQSKENQVSAKGTWRKIKENBS-----PTNISTQTSVSGATNGAESKTLIYOMA 2648
D 2112 PGT-----TWRIKQEDVPHILASTLPATLPLRVSSPEDSPAGTQ----- 2153
QY 2649 PAVSKTEDEVVRIEDCPINPRGRST--GNTPPVIDSVSEKANPNIKDKONQAKQNV 2706
D 2154 ---RKTSDAVQTEADVATSKTNSSTSPLESRDPP-----QAP--- 2188
QY 2707 GNGSVPMRTVGLNRLNSFTIOVDAPQKGTETKPGQNNPVVSETNESSIVERTPTSSSS 2766
D 2189 --ASGVAPQSGD-----VDGP-----VLITFPASAPP--HEGLSVAIRAGFP----- 2227
QY 2767 SSKHSPSGTVAARVTPFNPNPSPRKSS--ADSTSPASQPIPTPVN 2810
D 2228 TSHGSPSR--AARVPPFNTPVSPMAAATMAASDAVEKAFVSPAS 2271

```

```

RESULT 12
US-10-267-502-371
; Sequence 371, Application US/10267502
; Publication No. US20040071700A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jaeseob
; TITLE OF INVENTION: Obesity Linked Genes
; FILE REFERENCE: LSD-07416
; CURRENT APPLICATION NUMBER: US/10/267,502
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 371
; LENGTH: 2303
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-267-502-371

Query Match      23.4%; Score 3415.5; DB 12; Length 2303;
Best Local Similarity 34.4%; Pred. No. 8.4e-184;
Matches 997; Conservative 372; Mismatches 815; Indels 717; Gaps 94;

QY      2 AAASYDQLLKQVEALKVENSNLRQLEEDNSNHLTKLETEASNMKVEVLKQCGSIEDA-- 59
D 4 SVAPYEQLVRQVEALKAEHSLRQELADNSHLSKLETETSGMEVLKHQGLQKEARV 63
QY 60 MASSQIDLLRLKELNLDSSNFFGVKLRKMSLURSTGSRGSGSVSSGECSPVPMGSPF 119
D 64 LVSSQTEVLQKALQMDITSLYNLKFPDP-----TLGPPEP---AARTPEGSPV-HGSGP 115
QY 120 RR-GPVNCSRESTGVLEBELEKERSILLADIDKEKEKDYVYAOQLNLTKRIDSPLTE-N 177
D 116 SKDSFGLSRATIRLLBELDRERFLNEIEKEKEKLYWYQGLSKLDELPHVETQ 175
QY 178 FSLQDMMTRRLQVEARQIRVAMEEQIGTQODMBKRAQRRIARIQOIEKIDILRRQLQS 237
D 176 FSMQMDLIRQLEFEAQHRSIMEERFSGTDEMYQRAIRASLEQIDKELLE----- 228
QY 238 QATEARSSQNKHETGSHDAERQNGGCVGCHINWATSGNGSGSTTRMDHETASVLSST 297
D 229 -----AODRVQQTPEQALLAV-----KSPVDEDEPETEVPTHE 262
QY 298 KSAPRLTSLHGTQVEMVYSLLSMLGTHDKDDMRITLLAMSSSQDSCISMRQSGCLPLLI 357
D 263 DGTPOPGNS---KVEVVFWLLSLMTRDQEDTARTLLAMSSSPESCVMARSGCLPLLI 318
QY 358 QLLHGNDKDS---VLLGNSRSGSKARASAAHNIHISQPDDEKRGREIRVHLHQR 414
D 319 QILHTEAAGGRAGAPGAPGAKOARMRANAALHNIIVFSQPDQGLARKEKMKVHLVQR 378
QY 415 AYCETCWEQEAHPGMD-QDKNPMAPVEHQICPAVCVLMKLSFDEBHRHANNELGQL 473
D 379 AYCETCDMLQARDGGPEGGGAGSAPIEPQICQATCAVMKLSFDEEYREANMELGQL 438
QY 474 ATAEQLQVDCMYGLTNDHYITLRRYAGMALTNLTFGDVANKATLCSMKGCMALVAQL 533
D 439 AVAEQLQVDYEMHKMTRDPLNLRVYAGMPLTNLTFGDVANKATLCARRGCMALVAQL 498
QY 534 KSESDLOQVLTASVLNLSNRADVNSKKTLEVGSKVLMCALEVKESTKLSVLSALW 593
D 499 ASDSEELQVSSILNLSNRADINSKVLREASVTALVQCULRATKSTLKSLSALW 558
QY 594 NLSAHTENKADICAVDQALAFVLTGLTYRSQTNLTALIESGGGILNRVNSLIATNDHR 653
D 559 NLSAHTENKAAICQVQDGLGFLVSTLTLYKQCSNLAIIESGGGILNRVNSLVATREDYR 618
QY 654 QILRENNCLQTLLOHLKSHSLTIVSNACGTLWNLSARNPKQDEALWDMGVSMKLNLIHS 713
D 619 QVLRDHNCLQTLLOHLTSHSLTIVSNACGTLWNLSARSDQELLWDLGAVGMRLNVHS 678
QY 714 KHKMTAMGSAALRNLMANRPAYK-DANIMSPGSSLPFSLHVRKQKALEAELDAQHSET 772

```





RESULT 13  
 US-10-267-502-368  
 ; Sequence 368, Application US/10267502  
 ; Publication No. US20040071700A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kim, Jaeseob  
 ; TITLE OF INVENTION: Obesity Linked Genes  
 ; FILE REFERENCE: LSD-07416  
 ; CURRENT APPLICATION NUMBER: US/10/267,502  
 ; CURRENT FILING DATE: 2003-01-27  
 ; NUMBER OF SEQ ID NOS: 439  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 368  
 ; LENGTH: 2417  
 ; TYPE: PRT  
 ; ORGANISM: Drosophila melanogaster  
 US-10-267-502-368  
 Query Match 12.1%; Score 1765; DB 12; Length 2417;  
 Best Local Similarity 25.3%; Pred. No. 2.2e-90;  
 Matches 719; Conservative 325; Mismatches 838; Indels 964; Gaps 100;  
 QY 243 ERSQKHETGSHDAERON-EGQGVGEINMATSGGQSTTR-----MDHETASVLSSS 295  
 DB 129 ELREMRHSLDRNFERQSAQQQDLDELPFRNGGSPASAGRPSRSGKPSYTLRSLFDGD 188  
 QY 296 STHSAPR-----RLTSH-----LGTKEVMVYSLLSMLGTHDKDDMSRTLLAM 337  
 DB 189 APAPPLPKGAAWTTSFDERYTSVAETLGSKEVCVYSLLSMLGNSDPLEWAKKFLLEL 248  
 QY 338 SSGODSCISMRQGCLPLLIQLHGNKDSVLLGNRSGSKAPARASAAALHNIHSQPD 397  
 DB 249 SGNAQSCATLRRSGCMPLLVQWHPDND-----QEVKRCAGQALHNVVHSHDPE 298  
 QY 398 KRGRREIRVHLLEQIRAYCETCWENQEAHEPGM--DQDKNMPAPVHEQICPACVLMK 455  
 DB 299 KAGREAKVRLDQIVDYCSFLKTLQSGGEAIADDSRHL-----AAISLMK 349  
 QY 456 LSFDEEHRHAMNELGGLOAIALLQVDCENMYG-LTNDHYSITLRRYAGMALTNLTFGDVA 514  
 DB 350 VSFDEEHRHAMCELGAHAIPLNVLHDHVAHVGKPEDQCCNSLRRYALMALTNLTFGDEN 409  
 QY 515 NKATLSMKGCMALVAQLKSESDLOQVATSVLRNLSHRADVNSKTLREVGSVKALME 574  
 DB 410 NKALLCQKQFMALVAQLDSAPDDLQVTVASVLRNLSWRADSNMKAVLNEIGTVTALAL 469  
 QY 575 CALEVKKESTLKSVLNALMSLHACTENKADICAVDGAFLVGLTYRSTQNTLAIIES 634  
 DB 470 AAMNRSNTLKALLSALNLSAHCSTNKAEPICAVDGAFLVGLMSYEGPSKTLKIEN 529  
 QY 635 GGGILRVSSLIATNEDHRLRENKCLQTLQHLKSHSLITIVSNACGTLMLNLSARNPKD 694  
 DB 530 AGGILRVSSHIAVCEPYQILRHQNCALTLQQLKSESLITVVSNSCGTLNLSARSAD 589  
 QY 695 QEALWDGAVSMKLNLIHSKHKIAMSAAALNMANRPAYK-----DANIMSPG-SSL 749  
 DB 590 QKFLWDNGAVPMRLSLIHSKHAMISEGSSALKNLNFPAVQNHQLDPIARSMGLKAL 649  
 QY 750 PSLVFRKQKALEAEALDAQHLSEFTDNIDNLSPKASHRSKQRHKQSLGYDVFDTNRHDDN 809  
 DB 650 PTLKARKAKALQELGHRHTAETCDNL----- 677  
 QY 810 RSDNFNTGNVTLSPYNTTIVLSPSSSSRGLSDSRSEKDRSLERERGLGNHYHPATEN 869  
 DB 678 -----TGG-----KLDKERASSSSRRHP----- 695  
 QY 870 PGTSSKGLQISTTAAQIAKAMEVEVAIHTSQDRSSGSGTTELHCVTDBERNALRRSSAAH 929  
 DB 696 -----APRLTR-----SAMLTKSSRDSVTSKSDCAVDH----- 725

QY 930 THSNTYNTFTKSGNSRNTCTMPYAKLEYKRSNSDNLNVSSSDGYGKRGOMKPSIESYSED 989  
 DB 726 -----LIRSASASDAHRK-----VKPKITDFDLE 749  
 QY 990 DESKFCYSQYPADLAHKIHSANHMNDNDGELDTPTINYSLKYSDLEQNSGRQSPQNERW 1049  
 DB 750 ME-----QDTEATEEQPIDYSVKYS----- 769  
 QY 1050 ASPKHIIIEDEIKQSEQRORNRQSTTYPVYTESTDDKHLKQFHPFGQCECVSPYRSRGANG 1109  
 DB 770 -----ENATKISTY----- 778  
 QY 1110 SETNRVSGNHGINQVNSQSCQEDDYEDDKPTNYSERYSEEEHBEERPTNYSIKYNEE 1169  
 DB 779 -----QETDL--DQPTDFSLYAE-----NQIESDLDISGPAGGQ 811  
 QY 1170 KEHVDPIDYSLKYATDIP--SSQKQSFSSKSSSQSQSKTEH-----MSSSENSTSTPSS 1223  
 DB 812 KSTITPP-----AETVPEKSEGEILLILDDSVKCYQETDTPYVISNAASVTDLRVAA 864  
 QY 1224 NAKRQNLHPS--SAQSRSGQFQ--KAATCKVASINQETIQTICYVEDTPICFRCSSLS 1278  
 DB 865 KADAEAEVKEPREVRETSKEGAPKXLPKLSQCGSGSYTPKEPINYCEBGTGPGYFSKYDLS 924  
 QY 1279 SLSSAEDBIG--CNQTTQEDASANTLQIAEIKKIGTRSAEDPVSEVPAVSOHPKTSRSL 1337  
 DB 925 SL-----DESGKANAQIVGTD-----ADIKPKLEKQEBQESQPAEQVLTKPPTQANSAL 973  
 QY 1338 QGSSLSSESARHKAVERFSGAKSPSKGAQTPKPPPEHYVQETPLMFSTRCTVSSSLDSFE 1397  
 DB 974 -----ETPLMRSRSMDSLVHDP 992  
 QY 1398 SRSTA-----SSVQSEPCSGMVSGIISPSDLDPSPGQTPMPSRSKTPPPPPQTAQTKREV 1452  
 DB 993 DVDVANCDDKSSVVD--FSRLASGVSFSEIPDPSFTQSPMQS----- 1033  
 QY 1453 PKNKAPTAERESGPKQAAVNAQVQVLPDADTLHFAETESTPDGFCSSSLSALS-- 1511  
 DB 1034 PRNSVAGSGNVDSPPVVPASLOPLRSVPE--DDLSSFVVEHTPAQFSTATSLNSLIV 1092  
 QY 1512 -DEPFIQKQVELRIMPVQENDNGNE-----TESEQKSENQEKAE-AEKT 1556  
 DB 1093 DDE-----KAPASVAEDNEDELLANCINMGQRKPTFAVKSTVNVSEVDVAEET 1143  
 QY 1557 IDS-----EKD-----LLDD--SDDDDI-- 1572  
 DB 1144 IRSYCTEDTPALLSKVPSNTNLSVISMSTDPDKDATAGQAQMAHQSDVDVSSNASDCGG 1203  
 QY 1573 --EILEECIISAMPTKSSRKAKKPAQTAGKLPPPVARK-PSQLPVYKLLPSONRLQPK 1628  
 DB 1204 ASGHLQOQIRDGM-----KKPLGEATSDPTAMLRGNGNELPGY--LPS----- 1245  
 QY 1629 HVSTPGDDMPRVYCVGEPTINPSTATSLDLTIES-----PPNELAAGEVGRGAQSGEP 1684  
 DB 1246 -----ADEMNK-FLVEDSPCNFVSWGLNLTVGSSLVGPAVOLKETE--PSSADQNPEN 1297  
 QY 1685 EKRTIPTEGRSTDEAQQGKTSSVTIPELDDNKAEBGDILAEC----- 1727  
 DB 1298 KAKPGKQOVRPFWQDDSLSLSDSEDDTNLLSQAIANGCNRPKSNLGFSSNGKSS 1357  
 QY 1728 -----INSAMPKSKHKPFRVKIMQVQAQASASSAPNKNQ-----LDG-- 1767  
 DB 1358 SLSSSQPIAINAATASGLSASMTVRKSQQQESYSSVDSDSDNDNQSKSLFELCILKGMV 1417  
 QY 1768 KKKXP-----TSPVKPIQNTTEYRT-----RVRKNADSKNLNA 1801  
 DB 1418 KTKEPGARAQVQOQPIVGSVSSVQSNFSLKQFDSLVLQPSGGGVKQORHHHHHHRER 1477  
 QY 1802 ERVFSNDKDSK-KQNLKKN--SKDFNDKLPNNEDRVGSPAFSPHHYTFIEGTPYCFSR 1858  
 DB 1478 ER--ERRKDEKLLQECINTGTSKKIN-AVPEKNV--LATSAALPEPCH--PMAAT----- 1523  
 QY 1859 NDSLSSLDFFDDDDVDLSREKAEALPKAKENKESAKVTSHTELTSNQOSANKTQAIKQPI 1918





QY	2702	AKQNVGNGSVPMRTVTVGLENLNSFIQVDAPDQKGTETKPGQNNPVPVSETNESSIVERTP	2761
Db	1566	AKAPI---SAPFVHEGLGVAVGGF-----	1586
QY	2762	FSSSSSKHSGSPGCTVAAAVTPEYNNPSPRKSSADSTSA	2800
Db	1587	-----PASRHGSPSR--SARVPPFNYVPSMVAATTDSA	1619

Search completed: August 25, 2004, 17:53:02  
Job time : 215.5 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2004, 17:08:42 ; Search time 37.5 seconds

(without alignments)  
3913.937 Million cell updates/sec

Title: US-09-442-489F-7

Perfect score: 14566

Sequence: 1 MAAASYDQLLKQVEALKMEN.....ESSGTQSPKRGSGSYLTVS 2843

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA.\*
- 1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	14566	100.0	2843	1	US-08-452-655B-2
2	14566	100.0	2843	1	US-08-452-655B-7
3	14566	100.0	2843	3	US-08-450-582-2
4	14566	100.0	2843	3	US-08-450-582-7
5	14566	100.0	2843	2	US-08-821-355A-7
6	14566	100.0	2843	2	US-09-003-687A-7
7	14566	100.0	2843	3	US-09-136-605-7
8	14548.5	99.9	2842	1	US-07-741-940-7
9	14548.5	99.9	2842	1	US-08-289-548A-7
10	14548.5	99.9	2842	4	US-08-452-654-7
11	14548.5	99.9	2842	4	US-08-449-731-7
12	14533	99.8	2843	1	US-07-741-940-2
13	14533	99.8	2843	1	US-08-289-548A-2
14	14533	99.8	2843	1	US-08-452-654-2
15	14533	99.8	2843	2	US-08-370-235A-2
16	14533	99.8	2843	4	US-08-449-731-2
17	507.5	3.5	178	2	US-08-630-822A-100
18	507.5	3.5	178	2	US-09-005-063-100
19	507.5	3.5	178	4	US-09-171-156A-49
20	507.5	3.5	178	4	US-09-004-730A-49
21	507.5	3.5	178	4	US-08-981-799A-49
22	502.5	3.4	2137	4	US-09-134-001C-4463
23	461.5	3.2	10182	4	US-09-134-001C-3159
24	413	2.8	3596	4	US-09-134-001C-5080
25	385	2.6	8991	4	US-08-714-741-32
26	382.5	2.6	3256	4	US-09-519-172-98
27	382.5	2.6	3256	4	US-09-976-594-22

28	380.5	2.6	3248	1	US-08-353-700-1	Sequence 1, Appli
29	380.5	2.6	3248	5	PCT-US95-16216-1	Sequence 1, Appli
30	377.5	2.5	1581	3	US-09-110-517-2	Sequence 2, Appli
31	370.5	2.5	3969	3	US-08-061-376-5	Sequence 5, Appli
32	363	2.5	2468	4	US-09-976-594-726	Sequence 726, App
33	361	2.5	2662	4	US-09-595-684B-31	Sequence 31, Appli
34	346	2.4	2954	4	US-09-150-867-1	Sequence 1, Appli
35	337.5	2.3	2185	4	US-09-854-856-36	Sequence 36, Appli
36	337.5	2.3	2245	4	US-09-854-856-4	Sequence 4, Appli
37	337.5	2.3	2322	4	US-09-854-856-34	Sequence 34, Appli
38	337.5	2.3	2382	4	US-09-854-856-2	Sequence 2, Appli
39	332.5	2.3	2482	1	US-08-328-254-6	Sequence 6, Appli
40	322	2.2	2157	4	US-09-854-856-52	Sequence 52, Appli
41	322	2.2	2217	4	US-09-854-856-20	Sequence 20, Appli
42	322	2.2	2294	4	US-09-854-856-50	Sequence 50, Appli
43	322	2.2	2354	4	US-09-854-856-18	Sequence 18, Appli
44	313.5	2.2	1596	4	US-08-978-277A-4	Sequence 4, Appli
45	308	2.1	1969	4	US-09-418-710-72	Sequence 72, Appli

ALIGNMENTS

RESULT 1  
US-08-452-655B-2  
; Sequence 2, Application US/08452655B  
; Patent No. 5783666  
; GENERAL INFORMATION:  
; APPLICANT: ALBERTSEN, HANS  
; APPLICANT: ANAND, RAKESH  
; APPLICANT: CARLSON, MARY  
; APPLICANT: GRODEN, JOANNA  
; APPLICANT: HEDGE, PHILIP J.  
; APPLICANT: JOSLYN, GEOFF  
; APPLICANT: KINZLER, KENNETH  
; APPLICANT: MARKHAM, ALEXANDER F.  
; APPLICANT: NAKAMURA, YUSUKE  
; APPLICANT: THLIVERIS, ANDREW  
; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC  
; TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS  
; NUMBER OF SEQUENCES: 102  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Banner & Witcoff, Ltd.  
; STREET: 1001 G Street, NW  
; CITY: Washington  
; STATE: D.C.  
; ZIP: 20001-4598  
; COUNTRY: USA  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/452,655B  
; FILING DATE: 25-MAY-1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/289,548  
; FILING DATE: 12-AUG-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/741,940  
; FILING DATE: 08-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kagan, Sarah A.  
; REGISTRATION NUMBER: 32,141  
; REFERENCE/DOCKET NUMBER: 1107.49964  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-508-9100  
; TELEFAX: 202-508-9299  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2843 amino acids



TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-452-655B-2

Query Match 100.0%; Score 14566; DB 1; Length 2843;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2843; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MAAASVDQLLKQVEALQWENLNRLQLEEDNSNHLTKLETEASNKKEVLKQLGSIIDEAM	60
Db	1	MAAASVDQLLKQVEALQWENLNRLQLEEDNSNHLTKLETEASNKKEVLKQLGSIIDEAM	60
Qy	61	ASSGQIDLLRLKELNLDSSNFFGVKLRKMSLRYSRGSGVSSRGSCSPVPMGSPFR	120
Db	61	ASSGQIDLLRLKELNLDSSNFFGVKLRKMSLRYSRGSGVSSRGSCSPVPMGSPFR	120
Qy	121	RGFVNGSRESTGYLEELKERSLILLADLDKEEKEKDWYQAQLONLTKRIDSLPTENFSL	180
Db	121	RGFVNGSRESTGYLEELKERSLILLADLDKEEKEKDWYQAQLONLTKRIDSLPTENFSL	180
Qy	181	QDTMTRSQLYEAFQIRVAMEEQLGTQDMKEKAQRRIARIQIQIEKDILIRIQLLOSQAT	240
Db	181	QDTMTRSQLYEAFQIRVAMEEQLGTQDMKEKAQRRIARIQIQIEKDILIRIQLLOSQAT	240
Qy	241	EAERSSQNHKETSCHDAERQNEGGVGEINMATSGNGQSTTRMDHETASVSSSSSTHSA	300
Db	241	EAERSSQNHKETSCHDAERQNEGGVGEINMATSGNGQSTTRMDHETASVSSSSSTHSA	300
Qy	301	PRRLTSLHGTKEVMYVSLLSMLGTHDKDMSRTILLAMSSQDSCIWMROSGLPLLIQLL	360
Db	301	PRRLTSLHGTKEVMYVSLLSMLGTHDKDMSRTILLAMSSQDSCIWMROSGLPLLIQLL	360
Qy	361	HGNDKDSVILGNRSGSEKARASAAALHNIHISQPDQKGRREIRVLHLEQIRAYCETC	420
Db	361	HGNDKDSVILGNRSGSEKARASAAALHNIHISQPDQKGRREIRVLHLEQIRAYCETC	420
Qy	421	WEWQEAHEPGMDQKNPMPAPVPHOICPAVCVLMKLSFDEEHRHMANEELGGQAIAELQ	480
Db	421	WEWQEAHEPGMDQKNPMPAPVPHOICPAVCVLMKLSFDEEHRHMANEELGGQAIAELQ	480
Qy	481	VDCEMYGLTNDHYSITLRRYAGMALTNLTFDGVANRATILCSMKGMCRALVAQKSEDL	540
Db	481	VDCEMYGLTNDHYSITLRRYAGMALTNLTFDGVANRATILCSMKGMCRALVAQKSEDL	540
Qy	541	QOVTASVLRNLWRADVNSKKTIREVGSVKALMECALEVKKESTLKSVALWNLSAHT	600
Db	541	QOVTASVLRNLWRADVNSKKTIREVGSVKALMECALEVKKESTLKSVALWNLSAHT	600
Qy	601	ENKADI CAVDGALAFVLTGYTSQNTWLAIEBSGGGILRNVSLLIATNEDHRQILRENN	660
Db	601	ENKADI CAVDGALAFVLTGYTSQNTWLAIEBSGGGILRNVSLLIATNEDHRQILRENN	660
Qy	661	CLQTLQHLKSHSLITVSNACGLTNLWLSARNPKDQALWDMGAVSMLKNIHISKHNIAM	720
Db	661	CLQTLQHLKSHSLITVSNACGLTNLWLSARNPKDQALWDMGAVSMLKNIHISKHNIAM	720
Qy	721	GSAALRNLMANRPKADANIMSGSSLPVSLHVRKQKALEADLAHLSETFDNIDNLS	780
Db	721	GSAALRNLMANRPKADANIMSGSSLPVSLHVRKQKALEADLAHLSETFDNIDNLS	780
Qy	781	PKASHRSKQHKSLYGDYVDFDNRHDDNPNFNTGNMTVLSPLYNTTVLSPSSSSRGS	840
Db	781	PKASHRSKQHKSLYGDYVDFDNRHDDNPNFNTGNMTVLSPLYNTTVLSPSSSSRGS	840
Qy	841	LDSSRSKDRSLRERIGIGLVNHPATENPGTSSKRGLOISTTAAQIAKMEVEVAIHTS	900
Db	841	LDSSRSKDRSLRERIGIGLVNHPATENPGTSSKRGLOISTTAAQIAKMEVEVAIHTS	900
Qy	901	QEDRSSGSGTTELHCVTDERNALRRSAAHTHNTYNTFTKSENRTCSMPYAKLEYKRSS	960
Db	901	QEDRSSGSGTTELHCVTDERNALRRSAAHTHNTYNTFTKSENRTCSMPYAKLEYKRSS	960

Qy	961	NDLSNVSSSDGYGKRGQMKFPIESYSEDDDESKFCYSYGQYPADLAHKIHSANHMDNDGE	1020
Db	961	NDLSNVSSSDGYGKRGQMKFPIESYSEDDDESKFCYSYGQYPADLAHKIHSANHMDNDGE	1020
Qy	1021	LDTPINSLKYSDQLNSGRQSPQNERWARPPKHIIIEDEIKQSEORQSRQSTIYPVYTE	1080
Db	1021	LDTPINSLKYSDQLNSGRQSPQNERWARPPKHIIIEDEIKQSEORQSRQSTIYPVYTE	1080
Qy	1081	STDDKHLKFKFHFQGCQECVSPYRSGANGSETNVRGSHNGINQNVOSLQCOEDDYEDDKP	1140
Db	1081	STDDKHLKFKFHFQGCQECVSPYRSGANGSETNVRGSHNGINQNVOSLQCOEDDYEDDKP	1140
Qy	1141	TNYSERVSEEBQHEEBEPPTYSIKYNEEKHVDPIDYSLKYATDIPSSQKQSFSSKS	1200
Db	1141	TNYSERVSEEBQHEEBEPPTYSIKYNEEKHVDPIDYSLKYATDIPSSQKQSFSSKS	1200
Qy	1201	SSGSSKTEHMSSESSENTSPSSNAKQNLHPSSAQSRGQPOKAATCKVSSINQETIQ	1260
Db	1201	SSGSSKTEHMSSESSENTSPSSNAKQNLHPSSAQSRGQPOKAATCKVSSINQETIQ	1260
Qy	1261	TYCVEDTPICFSRCSLSIASSAEDTGCNQTTOEADSAANTLQIAETKEKIGTRSAEDPV	1320
Db	1261	TYCVEDTPICFSRCSLSIASSAEDTGCNQTTOEADSAANTLQIAETKEKIGTRSAEDPV	1320
Qy	1321	SEVPAVSOHPPTKSRRLQSGSLSSSESARHKAVERPSSGAKSPSKSGAOTPKSPPEHYVOET	1380
Db	1321	SEVPAVSOHPPTKSRRLQSGSLSSSESARHKAVERPSSGAKSPSKSGAOTPKSPPEHYVOET	1380
Qy	1381	PLMFSRCTSVSSLDSPESRSIASSVQSEPCSGMVSGIISPSDLPDPSQGTMPSPSRKTPP	1440
Db	1381	PLMFSRCTSVSSLDSPESRSIASSVQSEPCSGMVSGIISPSDLPDPSQGTMPSPSRKTPP	1440
Qy	1441	PPQTAQTKREVPKNAKPTAKESGPKQAANNAVQVLPDADTLLHFAFESTPDGF	1500
Db	1441	PPQTAQTKREVPKNAKPTAKESGPKQAANNAVQVLPDADTLLHFAFESTPDGF	1500
Qy	1501	SCSSLSALSDELPEFIQKVELRIMPVQVNDNGNETESQPKESNENQEKAKTIDSE	1560
Db	1501	SCSSLSALSDELPEFIQKVELRIMPVQVNDNGNETESQPKESNENQEKAKTIDSE	1560
Qy	1561	KDLDDDDDDDIIELEECIIISAMPTKSRKAKPAQATASKLPPPVARKPSQLPVYKLLPS	1620
Db	1561	KDLDDDDDDDIIELEECIIISAMPTKSRKAKPAQATASKLPPPVARKPSQLPVYKLLPS	1620
Qy	1621	QNRLOPKQHVSTFGDDMPRVYCVVEGTPINFSTATSLDITIESSPNELAAEGVRCQAQ	1680
Db	1621	QNRLOPKQHVSTFGDDMPRVYCVVEGTPINFSTATSLDITIESSPNELAAEGVRCQAQ	1680
Qy	1681	SGFEFEKRDITPTEGRSTDEAQQGKTSSTVITPELDNKAEBGDIILAEICINSAMPKGSHP	1740
Db	1681	SGFEFEKRDITPTEGRSTDEAQQGKTSSTVITPELDNKAEBGDIILAEICINSAMPKGSHP	1740
Qy	1741	FRVKIKIDVOQOASASSAPNKLQDKKKKPTSPVKPIQNTTEYRTRVRKNADSKNNLN	1800
Db	1741	FRVKIKIDVOQOASASSAPNKLQDKKKKPTSPVKPIQNTTEYRTRVRKNADSKNNLN	1800
Qy	1801	AERVFSDKNSKKQNLKNNKDFNDKLPNNEDRVGSAFDSPPHYTPIEGTYPYCFSRND	1860
Db	1801	AERVFSDKNSKKQNLKNNKDFNDKLPNNEDRVGSAFDSPPHYTPIEGTYPYCFSRND	1860
Qy	1861	SLSLDDDDDDVLSRKAELRKAKEKSEAKVTSHTELTSNQOSANKTQAKOPINR	1920
Db	1861	SLSLDDDDDDVLSRKAELRKAKEKSEAKVTSHTELTSNQOSANKTQAKOPINR	1920
Qy	1921	QCPKPIILQKOSTFPQSSKDIIPDRGAATDEKIQNPAIENTPVCFSHNSLSLSDIDQENN	1980
Db	1921	QCPKPIILQKOSTFPQSSKDIIPDRGAATDEKIQNPAIENTPVCFSHNSLSLSDIDQENN	1980
Qy	1981	KNENEPKETEPPDPSQESPKQASGVAPKSFHVEDTPVCFSRNSSLSSLSIASEDDLQ	2040
Db	1981	KNENEPKETEPPDPSQESPKQASGVAPKSFHVEDTPVCFSRNSSLSSLSIASEDDLQ	2040
Qy	2041	ECISAMPKPKKKPRLKGDNEKHSRPNMGILGDLTLDLKDICRDPDSEHGLSPDSENF	2100

Db 2041 ECISAMPKKKPSRLKGDNEKXSPRNMGGILGEDLTLDKDIQRDSEHGLSPDSENF 2100  
Qy 2101 WFAIOEGANSIVSSIHORAAAACLRQASSSDSDSLKSGISLSPHLLPDQOEKPT 2160  
Db 2101 WFAIOEGANSIVSSIHORAAAACLRQASSSDSDSLKSGISLSPHLLPDQOEKPT 2160  
Qy 2161 SNKGRILKPGKSTLETKKIBESKGIKGGKVKYSLITGKVRNSISGQMQLQAN 2220  
Db 2161 SNKGRILKPGKSTLETKKIBESKGIKGGKVKYSLITGKVRNSISGQMQLQAN 2220  
Qy 2221 MFSISGRMTIHPGVNRSSSTSVSKKGPPLKTPASKSPSEGTATTSRGAKPVS 2280  
Db 2221 MFSISGRMTIHPGVNRSSSTSVSKKGPPLKTPASKSPSEGTATTSRGAKPVS 2280  
Qy 2281 ELSPVARQTSQGGSKAPSRSGSDSTPSRPAQPLSRPQSPGRNISPGRNISP 2340  
Db 2281 ELSPVARQTSQGGSKAPSRSGSDSTPSRPAQPLSRPQSPGRNISPGRNISP 2340  
Qy 2341 KLSQLPRTSSPTASTKSGSGKMYTSPGRMSQONLTQTGLSKNASSIPRESASK 2400  
Db 2341 KLSQLPRTSSPTASTKSGSGKMYTSPGRMSQONLTQTGLSKNASSIPRESASK 2400  
Qy 2401 LNMNGNGANKKVELSRMSSTKSGSDSDRERPVLRQSTFIKEAPSLRKLLESA 2460  
Db 2401 LNMNGNGANKKVELSRMSSTKSGSDSDRERPVLRQSTFIKEAPSLRKLLESA 2460  
Qy 2461 SPESLSPSRPASPTRSQAQTPVLSPLPDMSLSTHSSVQAGWKLPPNLSPTIEYNDG 2520  
Db 2461 SPESLSPSRPASPTRSQAQTPVLSPLPDMSLSTHSSVQAGWKLPPNLSPTIEYNDG 2520  
Qy 2521 RPAKHDTARSHSPSRPLPINSRGTWKRHSKHSSSLPRVSTWRTGSSSILSASSES 2580  
Db 2521 RPAKHDTARSHSPSRPLPINSRGTWKRHSKHSSSLPRVSTWRTGSSSILSASSES 2580  
Qy 2581 SEKAKSEDEKHVNSISGTKQSKENOVSAKGTWRKIKENEFPSTNSTQTSVSGATNGA 2640  
Db 2581 SEKAKSEDEKHVNSISGTKQSKENOVSAKGTWRKIKENEFPSTNSTQTSVSGATNGA 2640  
Qy 2641 KTLIIQMAPAVSKTEDVWRIEDCPINNPRSGRSTGNTPTVIDSVSEKANPNKSKDN 2700  
Db 2641 KTLIIQMAPAVSKTEDVWRIEDCPINNPRSGRSTGNTPTVIDSVSEKANPNKSKDN 2700  
Qy 2701 QAKQVNGSVPMRTVGLNRLNSFIQVADPOKTEIKPGQNNPVVSETNESSIVERT 2760  
Db 2701 QAKQVNGSVPMRTVGLNRLNSFIQVADPOKTEIKPGQNNPVVSETNESSIVERT 2760  
Qy 2761 PFSSSSSKHSPSGTVAARVTTPNPNPSRKSADSTARSQIPTPVNNTKKRDSKT 2820  
Db 2761 PFSSSSSKHSPSGTVAARVTTPNPNPSRKSADSTARSQIPTPVNNTKKRDSKT 2820  
Qy 2821 DSTESSGTQSPKRHSGSYLVTSV 2843  
Db 2821 DSTESSGTQSPKRHSGSYLVTSV 2843

RESULT 2  
US-08-452-655B-7

Sequence 7, Application US/08452655B  
Patent No. 5783666

GENERAL INFORMATION:

APPLICANT: ALBERTSEN, HANS  
APPLICANT: ANAND, RAKESH  
APPLICANT: CARLSON, MARY  
APPLICANT: GRODEN, JOANNA  
APPLICANT: HEDGE, PHILIP J.  
APPLICANT: JOSLYN, GEOFF  
APPLICANT: KINZLER, KENNETH  
APPLICANT: MARKHAM, ALEXANDER F.  
APPLICANT: NAKAMURA, YUSUKE  
APPLICANT: THLIVERIS, ANDREW  
TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC  
TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS

NUMBER OF SEQUENCES: 102  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Banner & Witcoff, Ltd.  
STREET: 1001 G Street, NW  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20001-4598  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/452,655B  
FILING DATE: 25-MAY-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/289,548  
FILING DATE: 12-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/741,940  
FILING DATE: 08-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Kagan, Sarah A.  
REGISTRATION NUMBER: 32,141  
REFERENCE/DOCKET NUMBER: 1107.49964  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-508-9100  
TELEFAX: 202-508-9299  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2843 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHEICAL: YES  
ANTI-SENSE: NO  
US-08-452-655B-7

Query Match 100.0%; Score 14566; DB 1; Length 2843;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2843; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAAASYDQLLKQVEALKMENSRLRQELDENSHTLTKLETSANNKEVYLKQSGSIEDEAM 60  
Db 1 MAAASYDQLLKQVEALKMENSRLRQELDENSHTLTKLETSANNKEVYLKQSGSIEDEAM 60  
Qy 61 ASSQIDLLERLRLKELNLDSSNFFGVKLRKMSLRSGSREGSVSSRSGECSVPVMSFPR 120  
Db 61 ASSQIDLLERLRLKELNLDSSNFFGVKLRKMSLRSGSREGSVSSRSGECSVPVMSFPR 120  
Qy 121 RGFVNGSRESTGYLELEKERSILLADLDKEEKEKWYIAQLQNLTKRIDSLPTENFSL 180  
Db 121 RGFVNGSRESTGYLELEKERSILLADLDKEEKEKWYIAQLQNLTKRIDSLPTENFSL 180  
Qy 181 QDTMTREQLVEARQIRVAMEEQLGTCODMEKRAORRIARIQIIEKDLIRIQLLOQAT 240  
Db 181 QDTMTREQLVEARQIRVAMEEQLGTCODMEKRAORRIARIQIIEKDLIRIQLLOQAT 240  
Qy 241 EAERSQNKHETGSHDAERQNEGGVGEINMATSNGQGGSTTRMDHETASVLSSTHSA 300  
Db 241 EAERSQNKHETGSHDAERQNEGGVGEINMATSNGQGGSTTRMDHETASVLSSTHSA 300  
Qy 301 PRRLTSHLGTGVEMVYSLLSMLGTHDKDDMSRTLLAMSSQSDSCISWRQSGCLPLLIQL 360  
Db 301 PRRLTSHLGTGVEMVYSLLSMLGTHDKDDMSRTLLAMSSQSDSCISWRQSGCLPLLIQL 360  
Qy 361 HGNDKDSVLLGNRSGSKAARASAAALHNIHSQPDQKRRRIRVILHLEQIRAYCETC 420  
Db 361 HGNDKDSVLLGNRSGSKAARASAAALHNIHSQPDQKRRRIRVILHLEQIRAYCETC 420

QY 421 WEOEAHEPGWODKQNPAPVHQCIPAVCVLMKLSFOEHRHAMNELGGLQAIABLLQ 480  
Db 421 WEOEAHEPGWODKQNPAPVHQCIPAVCVLMKLSFDEHRHAMNELGGLQAIABLLQ 480  
QY 481 VDCEMYGLTNDHYSITLRRYAGMALTNLTFGDVANKATLCMSKGMCRALVAQLKSSIDL 540  
Db 481 VDCEMYGLTNDHYSITLRRYAGMALTNLTFGDVANKATLCMSKGMCRALVAQLKSSIDL 540  
QY 541 QOVIASVLRNLWRADVNSKKTIREVGSVKALMECALEVKKESTLKSVALNLSAHT 600  
Db 541 QOVIASVLRNLWRADVNSKKTIREVGSVKALMECALEVKKESTLKSVALNLSAHT 600  
QY 601 ENKADICAVDGAFLVGLTYRSQNTNLAIIESGGILRNVSLLIATNEDHRQILRENN 660  
Db 601 ENKADICAVDGAFLVGLTYRSQNTNLAIIESGGILRNVSLLIATNEDHRQILRENN 660  
QY 661 CLOTLLOHLKSHSLTIIVSNACGLTNLSARNPKDOEALWDMGAVSMLKNLIHSHKHMIAM 720  
Db 661 CLOTLLOHLKSHSLTIIVSNACGLTNLSARNPKDOEALWDMGAVSMLKNLIHSHKHMIAM 720  
QY 721 GSAALRNLMANRPAPKADANIMSPGSLPSLVRKQKALEAELDAQHLSSETFDNIDNLS 780  
Db 721 GSAALRNLMANRPAPKADANIMSPGSLPSLVRKQKALEAELDAQHLSSETFDNIDNLS 780  
QY 781 PKASHRSKQKHQSLYGDYVDFNTRHDNRSDNFTGNMTVLSPYLNTTVLPSSSSSRGS 840  
Db 781 PKASHRSKQKHQSLYGDYVDFNTRHDNRSDNFTGNMTVLSPYLNTTVLPSSSSSRGS 840  
QY 841 LDDSRSEKDSLEBERGIGLGNVHPATENPGTSKGLQISTTAAOIAKVMREVSIAHTS 900  
Db 841 LDDSRSEKDSLEBERGIGLGNVHPATENPGTSKGLQISTTAAOIAKVMREVSIAHTS 900  
QY 901 QEDRSSGSTTELHCVTDERNALRRSSAAHTSNFTKSENNRCTCMPYAKLEYKRSS 960  
Db 901 QEDRSSGSTTELHCVTDERNALRRSSAAHTSNFTKSENNRCTCMPYAKLEYKRSS 960  
QY 961 NDSLNSVSSDGYCKRQOMKPSIESYEDDESFCYSQYPADLAHKIHSANHMDNDGE 1020  
Db 961 NDSLNSVSSDGYCKRQOMKPSIESYEDDESFCYSQYPADLAHKIHSANHMDNDGE 1020  
QY 1021 LDTPIYSLKYSDQLNSGRQSPQNERWARPKHIIIEDEIKQSEQRNRQSTTYPVYTE 1080  
Db 1021 LDTPIYSLKYSDQLNSGRQSPQNERWARPKHIIIEDEIKQSEQRNRQSTTYPVYTE 1080  
QY 1081 STDDXHLKFPQHFQOQECVSPYRSRGANGSETNRVGSNHGINQVQSGLCOEDDYEDDKP 1140  
Db 1081 STDDXHLKFPQHFQOQECVSPYRSRGANGSETNRVGSNHGINQVQSGLCOEDDYEDDKP 1140  
QY 1141 TNYSERVSEBOHEBERPTNYSTKYNEERKHVDQIDYSLKYATDIPSSQKQSFSPKS 1200  
Db 1141 TNYSERVSEBOHEBERPTNYSLKYNEERKHVDQIDYSLKYATDIPSSQKQSFSPKS 1200  
QY 1201 SSGQSKTEHMSSESSTSTPSSNAKQNLHPSSAQSRSGQPKAATCKVSSINQETIQ 1260  
Db 1201 SSGQSKTEHMSSESSTSTPSSNAKQNLHPSSAQSRSGQPKAATCKVSSINQETIQ 1260  
QY 1261 TYCVEDTPICFSRCSLSLSAABDEIGCNOQTCEADSNLTQIAETKEKIGTRSAEDPV 1320  
Db 1261 TYCVEDTPICFSRCSLSLSAABDEIGCNOQTCEADSNLTQIAETKEKIGTRSAEDPV 1320  
QY 1321 SEVPAVQHPRTKSRRLQSSLSSESARHKAVFSSGAKSPSKGAQTPKSPPEHYVQET 1380  
Db 1321 SEVPAVQHPRTKSRRLQSSLSSESARHKAVFSSGAKSPSKGAQTPKSPPEHYVQET 1380  
QY 1381 PLMFSRCTSVSSLDSFESRIASVQSEPCSGVSGHIIISPSDLPDPCQTMPSPRSKTPP 1440  
Db 1381 PLMFSRCTSVSSLDSFESRIASVQSEPCSGVSGHIIISPSDLPDPCQTMPSPRSKTPP 1440  
QY 1441 PPTQATQKEVPKPKAPTAKEKESGPKQAANVAQVQVLPDADTLHLHFAESTPDGF 1500  
Db 1441 PPTQATQKEVPKPKAPTAKEKESGPKQAANVAQVQVLPDADTLHLHFAESTPDGF 1500  
QY 1501 SCSSLSALSIDEPFIQKDVLRIMPVQENDNGNETESEQPKESNENQKEAEKTIIDSE 1560

Db 1501 SCSSLSALSIDEPFIQKDVLRIMPVQENDNGNETESEQPKESNENQKEAEKTIIDSE 1560  
QY 1561 KOLLDDSDDDDIIELEBECIISAMPTKSRKAKKPAQATASKLPPPVARKPSQLPVYKLLPS 1620  
Db 1561 KOLLDDSDDDDIIELEBECIISAMPTKSRKAKKPAQATASKLPPPVARKPSQLPVYKLLPS 1620  
QY 1621 QNRLOPQKHVSFTPGDDMPRVYCVEGTPIFSTATSLDITIESPPNLAAGEVGRGAQ 1680  
Db 1621 QNRLOPQKHVSFTPGDDMPRVYCVEGTPIFSTATSLDITIESPPNLAAGEVGRGAQ 1680  
QY 1681 SGEFEKRTIIPTEGRSTDEAOGKTSVTTPELDDNKAEBGDIILAEICINSAMPKGSHPK 1740  
Db 1681 SGEFEKRTIIPTEGRSTDEAOGKTSVTTPELDDNKAEBGDIILAEICINSAMPKGSHPK 1740  
QY 1741 FRVKKIMDQVOOQASASSAPNKNQLDGKKKPTSPVKPIPONTYRTRVRKNADSKNIN 1800  
Db 1741 FRVKKIMDQVOOQASASSAPNKNQLDGKKKPTSPVKPIPONTYRTRVRKNADSKNIN 1800  
QY 1801 AERVFSNDKOSKQNLKXNSKDFNDKLPNNEDRVGSAFDSPHHYTPIEGTPYCFSRND 1860  
Db 1801 AERVFSNDKOSKQNLKXNSKDFNDKLPNNEDRVGSAFDSPHHYTPIEGTPYCFSRND 1860  
QY 1861 SLSLDFDDDDVDLSREKAEKAKENKESAKVTSHTELTSNOOSANKTOAIKQPINR 1920  
Db 1861 SLSLDFDDDDVDLSREKAEKAKENKESAKVTSHTELTSNOOSANKTOAIKQPINR 1920  
QY 1921 GQPKPILQKQSTFPQSSKDIPDRGAATDEKLNFAIENTPVCFSHNSLSLSLSDIDQENN 1980  
Db 1921 GQPKPILQKQSTFPQSSKDIPDRGAATDEKLNFAIENTPVCFSHNSLSLSLSDIDQENN 1980  
QY 1981 NKENPPIKETPPPOSQGEPPKQAGYAPKSFHVEDTPVCFSRNSLSLSLSDIDEDLLQ 2040  
Db 1981 NKENPPIKETPPPOSQGEPPKQAGYAPKSFHVEDTPVCFSRNSLSLSLSDIDEDLLQ 2040  
QY 2041 ECISAMPKKKPRLKGDNEKHSFRNMGGITLGBDLTLDLKDIORPDSEHGLSPDSNFD 2100  
Db 2041 ECISAMPKKKPRLKGDNEKHSFRNMGGITLGBDLTLDLKDIORPDSEHGLSPDSNFD 2100  
QY 2101 WKAIQEGANSIVSSLHQAAAAACLSRQASDSDSILSKSGISLGSFPFLTTPDOEKPF 2160  
Db 2101 WKAIQEGANSIVSSLHQAAAAACLSRQASDSDSILSKSGISLGSFPFLTTPDOEKPF 2160  
QY 2161 SNKGPRILLKPEKSTLETTKIESKGIKGGKVKYKSLITGKVRNSSEISQMKQPLQAN 2220  
Db 2161 SNKGPRILLKPEKSTLETTKIESKGIKGGKVKYKSLITGKVRNSSEISQMKQPLQAN 2220  
QY 2221 MFSISRGRTMIHIFGVNRNSSSTSPVSKGPPKTPASKSPSEGOATTPRGAAPSVKS 2280  
Db 2221 MFSISRGRTMIHIFGVNRNSSSTSPVSKGPPKTPASKSPSEGOATTPRGAAPSVKS 2280  
QY 2281 ELSVPARQTSOIGSSKAPSRSGRSDSTSPRPAQOPLSRPIQSPGRNSISPRNGISPPN 2340  
Db 2281 ELSVPARQTSOIGSSKAPSRSGRSDSTSPRPAQOPLSRPIQSPGRNSISPRNGISPPN 2340  
QY 2341 KLSQLPRTSSPSTASTKSSGSGKMSYTSRQMSQOQLTKQTGLSKNASSIPRSESASKG 2400  
Db 2341 KLSQLPRTSSPSTASTKSSGSGKMSYTSRQMSQOQLTKQTGLSKNASSIPRSESASKG 2400  
QY 2401 LNMNMNGANKVKVELSRMSSTKSSGSDRSEPRVLRQSTFTKEAPSPTLRKLRESA 2460  
Db 2401 LNMNMNGANKVKVELSRMSSTKSSGSDRSEPRVLRQSTFTKEAPSPTLRKLRESA 2460  
QY 2461 SPESLSPSSRPASPTRSOAQTPVLSPLPDMSLSTHSSVQAGWRKLPPLNLSPIEYNDG 2520  
Db 2461 SPESLSPSSRPASPTRSOAQTPVLSPLPDMSLSTHSSVQAGWRKLPPLNLSPIEYNDG 2520  
QY 2521 RPAKEDHIAHSHSESPSLPINRSGTWKRKSHKSSSLPRVSTWRTGSSSSILSASES 2580  
Db 2521 RPAKEDHIAHSHSESPSLPINRSGTWKRKSHKSSSLPRVSTWRTGSSSSILSASES 2580  
QY 2581 SEKAKSEDEKHNSTSGTKQSKENQVSAKGTWRKIKENEFSTNISTQOTVSSSGATNGAES 2640

Db 2581 SEKASEDEKHVNSISGTHQSKENQVSAKGTWRKTKENEFSPTNSTQTSVSGATNGAES 2640  
Qy 2641 KTLIYQMAPAVSKTEDVWVRIEDCPINNPRSGRSTGNTPTPVIDSSEKAMPNIKDSKDN 2700  
Db 2641 KTLIYQMAPAVSKTEDVWVRIEDCPINNPRSGRSTGNTPTPVIDSSEKAMPNIKDSKDN 2700  
Qy 2701 QAKQNVGSGVPMRTVGLNRLNSFIQVADPOKTEIKPGQNNPVPVSETNESSIVERT 2760  
Db 2701 QAKQNVGSGVPMRTVGLNRLNSFIQVADPOKTEIKPGQNNPVPVSETNESSIVERT 2760  
Qy 2761 PFSSSSSKHSSPSTGAARVTPFNPNPSPRKSADTSARPQIPTPVNNNTKKRDSKT 2820  
Db 2761 PFSSSSSKHSSPSTGAARVTPFNPNPSPRKSADTSARPQIPTPVNNNTKKRDSKT 2820  
Qy 2821 DSTSSSGTQSPXRHSGSVLVTSV 2843  
Db 2821 DSTSSSGTQSPXRHSGSVLVTSV 2843

RESULT 3  
US-08-450-582-2  
; Sequence 2, Application US/08450582  
; Patent No. 6114124  
; GENERAL INFORMATION:  
; APPLICANT: ALBERTSEN, HANS  
; APPLICANT: ANAND, RAKESH  
; APPLICANT: CARLSON, MARY  
; APPLICANT: GRODEN, JOANNA  
; APPLICANT: HEDGE, PHILIP J.  
; APPLICANT: JOSLYN, GEOFF  
; APPLICANT: KINZLER, KENNETH  
; APPLICANT: MARKHAM, ALEXANDER F.  
; APPLICANT: NAKAMURA, YUSUKE  
; APPLICANT: THLIVERIS, ANDREW  
; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC  
; TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS  
; NUMBER OF SEQUENCES: 102  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Banner & Witcoff, Ltd.  
; STREET: 1001 G Street, NW  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20001-4598  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/450,582  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/452,655  
; FILING DATE: 25-MAY-1995  
; APPLICATION NUMBER: US 08/289,548  
; FILING DATE: 12-AUG-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/741,940  
; FILING DATE: 08-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kagan, Sarah A.  
; REGISTRATION NUMBER: 32,141  
; REFERENCE/DOCKET NUMBER: 1107.49964  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-508-9100  
; TELEFAX: 202-508-9299  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2843 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear

; MOLECULE TYPE: protein  
US-08-450-582-2  
Query Match 100.0%; Score 14566; DB 3; Length 2843;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2843; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MAAASYDQLLKQVEALKQENSNLRQELDESNHLTKLETEASNKEVYLKQLOGSIEDEAM 60  
Db 1 MAAASYDQLLKQVEALKQENSNLRQELDESNHLTKLETEASNKEVYLKQLOGSIEDEAM 60  
Qy 61 ASSQIDILLRLKELNLDSSNFFGVKLRKWSLRSGVREGSVSSRSGSGSPVPMGSFPR 120  
Db 61 ASSQIDILLRLKELNLDSSNFFGVKLRKWSLRSGVREGSVSSRSGSGSPVPMGSFPR 120  
Qy 121 RGFVNGSRESTGYLELEKERSLLADLDKEEKEKQWYLAQLQNLTKRIDSLPLTENFSL 180  
Db 121 RGFVNGSRESTGYLELEKERSLLADLDKEEKEKQWYLAQLQNLTKRIDSLPLTENFSL 180  
Qy 181 QTMTRQLEVEARQIRVAVEEOLGTCDMEKQAQRRIARIQOIEKDILRIOLLOSOAT 240  
Db 181 QTMTRQLEVEARQIRVAVEEOLGTCDMEKQAQRRIARIQOIEKDILRIOLLOSOAT 240  
Qy 241 EAERSSQNKHETGSHDAERQNEGQGVGEINMATSGNGQGSTTRMDHETASVLSSTHSA 300  
Db 241 EAERSSQNKHETGSHDAERQNEGQGVGEINMATSGNGQGSTTRMDHETASVLSSTHSA 300  
Qy 301 PRLTSHLGTKEVMYISLLSMGLTHDKDDMSRILLAMSSQDSCI SMRQSGCPLLIQLL 360  
Db 301 PRLTSHLGTKEVMYISLLSMGLTHDKDDMSRILLAMSSQDSCI SMRQSGCPLLIQLL 360  
Qy 361 HGNDKDSVLLGNRSGSKEARASAAALHNIHSOPDDKRGRRIRVHLLEQIRAYCETC 420  
Db 361 HGNDKDSVLLGNRSGSKEARASAAALHNIHSOPDDKRGRRIRVHLLEQIRAYCETC 420  
Qy 421 WEMQEAHEPGMDOKNMPAPVEHQICPAVCVLMKJSDFEHRHAMNELGGLQAIABLLQ 480  
Db 421 WEMQEAHEPGMDOKNMPAPVEHQICPAVCVLMKJSDFEHRHAMNELGGLQAIABLLQ 480  
Qy 481 VDCEMYGLTNDHYSITLRRYAGWALTNLTFGDVANKATLCMSKGMCPALVAQLKSEEDL 540  
Db 481 VDCEMYGLTNDHYSITLRRYAGWALTNLTFGDVANKATLCMSKGMCPALVAQLKSEEDL 540  
Qy 541 QQVIASVLRNLSWRADVNSKTLREVSGVKALMECALVKKESTLKSVALNLSAHCT 600  
Db 541 QQVIASVLRNLSWRADVNSKTLREVSGVKALMECALVKKESTLKSVALNLSAHCT 600  
Qy 601 ENKADICAVDGAFLVGTLYTSQNTNLTALIESGGIIRNVSSLIATNEDHQLRENN 660  
Db 601 ENKADICAVDGAFLVGTLYTSQNTNLTALIESGGIIRNVSSLIATNEDHQLRENN 660  
Qy 661 CLQTLQLKSHSLTI VSNACGTLNLSARNPKQDEALWDMGAVSMLKNIHSGKHMIAM 720  
Db 661 CLQTLQLKSHSLTI VSNACGTLNLSARNPKQDEALWDMGAVSMLKNIHSGKHMIAM 720  
Qy 721 GSAALNLMANRPAPKYKANDANIMSPGSSLSLHVVRKOKALEABLDQAHLSETFDNIDLS 780  
Db 721 GSAALNLMANRPAPKYKANDANIMSPGSSLSLHVVRKOKALEABLDQAHLSETFDNIDLS 780  
Qy 781 PKASHRSKORHKQSLYGDYVFDNRHDDNDSNFTGNMTVLSPLYNTTVLPSSSSSRGS 840  
Db 781 PKASHRSKORHKQSLYGDYVFDNRHDDNDSNFTGNMTVLSPLYNTTVLPSSSSSRGS 840  
Qy 841 LDSSRSKDRSLBERGIGLGNYPATENPGTSSKRLQISTTAAQIAKVMEEVSAIHTS 900  
Db 841 LDSSRSKDRSLBERGIGLGNYPATENPGTSSKRLQISTTAAQIAKVMEEVSAIHTS 900  
Qy 901 QEDRSSGTTTELHCVTDERNALRRSSAAHTSHNTYFTKSENGNRCTCSMPYAKLEYKRSS 960  
Db 901 QEDRSSGTTTELHCVTDERNALRRSSAAHTSHNTYFTKSENGNRCTCSMPYAKLEYKRSS 960  
Qy 961 NDSLNSVSSSDGYGKRGOMKPSIESYEDDESFKPSYGOYPADLAHKIHSANHMDNDGE 1020

Db 961 NDSLNSVSSDGYCKRQMKPSTESYEDDESKECSYGOYPADLAHKIHSANHWDDNDGE 1020  
Qy 1021 LDTPINTSLKYSDQLNGRQSPQNERWARPKHIIIEDEIKQSEQRNQSTTYPVYTE 1080  
Db 1021 LDTPINTSLKYSDQLNGRQSPQNERWARPKHIIIEDEIKQSEQRNQSTTYPVYTE 1080  
Qy 1081 STDDKHLKPHFQOQCVSPYSRGANGSETNRVGSNHGINONVSQSLCOEDDYEDDKP 1140  
Db 1081 STDDKHLKPHFQOQCVSPYSRGANGSETNRVGSNHGINONVSQSLCOEDDYEDDKP 1140  
Qy 1141 TNSERYSEOEHEBERPTNYSIKYNEEKHRVDQPIDYSLKYATDIPSSQKQSFSSKS 1200  
Db 1141 TNSERYSEOEHEBERPTNYSIKYNEEKHRVDQPIDYSLKYATDIPSSQKQSFSSKS 1200  
Qy 1201 SSGOSSKTEHMSSESSTSTPSSNAKRONOLHPSSAQSRGQPKAATCKVSSINQETIQ 1260  
Db 1201 SSGOSSKTEHMSSESSTSTPSSNAKRONOLHPSSAQSRGQPKAATCKVSSINQETIQ 1260  
Qy 1261 TYCVEDTPICFSRCSLSLSAEDDEIGCQNTTQEAADSANTLOIAETKEKIGTRSAEDPV 1320  
Db 1261 TYCVEDTPICFSRCSLSLSAEDDEIGCQNTTQEAADSANTLOIAETKEKIGTRSAEDPV 1320  
Qy 1321 SEVPAVQCHPRTKSRLQGSLSSESARHKAFFSSGAKSPSKGAGQTPKSPPEHYVQET 1380  
Db 1321 SEVPAVQCHPRTKSRLQGSLSSESARHKAFFSSGAKSPSKGAGQTPKSPPEHYVQET 1380  
Qy 1381 PLMFSRCTSVSSLDSPFSRSTASSVQSEPCSGVSGIISPSDLPDPGQTMPPSRKTPP 1440  
Db 1381 PLMFSRCTSVSSLDSPFSRSTASSVQSEPCSGVSGIISPSDLPDPGQTMPPSRKTPP 1440  
Qy 1441 PPPQTAQTKREVVPKNKAPTAKRESGPKQAANVAORVQVLPDADTLHFATESTPDGF 1500  
Db 1441 PPPQTAQTKREVVPKNKAPTAKRESGPKQAANVAORVQVLPDADTLHFATESTPDGF 1500  
Qy 1501 SCSSLSALSLDEPFIQKVELRIMPVQENDNGNETSEOPKESNENKEAEKIIDSE 1560  
Db 1501 SCSSLSALSLDEPFIQKVELRIMPVQENDNGNETSEOPKESNENKEAEKIIDSE 1560  
Qy 1561 KDLLDDDDDDIELEECIISAMPTKSSRKAKKPAQTAASKLPPPVARKPSQLPVYKLLPS 1620  
Db 1561 KDLLDDDDDDIELEECIISAMPTKSSRKAKKPAQTAASKLPPPVARKPSQLPVYKLLPS 1620  
Qy 1621 QNRLOPKHVSFTGDDMPRYVCEGTPIINFSTATSLSDLTIESPPNELAAGVGVRGAQ 1680  
Db 1621 QNRLOPKHVSFTGDDMPRYVCEGTPIINFSTATSLSDLTIESPPNELAAGVGVRGAQ 1680  
Qy 1681 SGEFEKDTTPTGSRSTDEAQGGTSSVTIPELDDNKAEBGDIILAEICINSAMPKGSHKP 1740  
Db 1681 SGEFEKDTTPTGSRSTDEAQGGTSSVTIPELDDNKAEBGDIILAEICINSAMPKGSHKP 1740  
Qy 1741 FRVKIMQVQOQASASSANQKQDGGKKKPTSPVKPIPQNTYTRVRKNADSKNNLN 1800  
Db 1741 FRVKIMQVQOQASASSANQKQDGGKKKPTSPVKPIPQNTYTRVRKNADSKNNLN 1800  
Qy 1801 AERFESDNKSKONLKNNSKDFNDKLPNNEDVRGSAFDSPHHYTPIEGTFCFSRND 1860  
Db 1801 AERFESDNKSKONLKNNSKDFNDKLPNNEDVRGSAFDSPHHYTPIEGTFCFSRND 1860  
Qy 1861 SLSSLDLDDDDVLSREKAEELKAKENKESAKVTSHTELTSNQOQANKTQAIAKQPINR 1920  
Db 1861 SLSSLDLDDDDVLSREKAEELKAKENKESAKVTSHTELTSNQOQANKTQAIAKQPINR 1920  
Qy 1921 GQPKPILQKOSTPQSSKQIDPDRGAATDEKLOFALIENTPVCFSHNSLSLSDIDQENN 1980  
Db 1921 GQPKPILQKOSTPQSSKQIDPDRGAATDEKLOFALIENTPVCFSHNSLSLSDIDQENN 1980  
Qy 1981 NKNEPIKETEPDPSQGEPSKQASGVAPKS FHVEDTFCVCFSRNSSLSLSDIDEDLLQ 2040  
Db 1981 NKNEPIKETEPDPSQGEPSKQASGVAPKS FHVEDTFCVCFSRNSSLSLSDIDEDLLQ 2040  
Qy 2041 ECISAMPKPKKPSRLKGDNEKSPRNMGILGEDTLIDLKOTORDSHGLSPDSNFED 2100  
Db 2041 ECISAMPKPKKPSRLKGDNEKSPRNMGILGEDTLIDLKOTORDSHGLSPDSNFED 2100

Qy 2101 WKATOEGANSIVSSILHQAAAACLSROASSSDSDSILSLKSGISLGSPFHLTPDOEEKPFT 2160  
Db 2101 WKATOEGANSIVSSILHQAAAACLSROASSSDSDSILSLKSGISLGSPFHLTPDOEEKPFT 2160  
Qy 2161 SNKGPRILLKPEKSTLETETKIIESSKGIKGGKVVYKSLITGKVSNSSEISQMKQPLQAN 2220  
Db 2161 SNKGPRILLKPEKSTLETETKIIESSKGIKGGKVVYKSLITGKVSNSSEISQMKQPLQAN 2220  
Qy 2221 MPSISRGRTMIHICGVNRNSSSTSPVKKGPPLKTPASKSPSECOQTATTSPRGAKPSVKS 2280  
Db 2221 MPSISRGRTMIHICGVNRNSSSTSPVKKGPPLKTPASKSPSECOQTATTSPRGAKPSVKS 2280  
Qy 2281 ELSVPARTSOIGGSSKAPSRSGSRDSTPSPAPQPLSRPIQSPGRNSISPGRNGISPPN 2340  
Db 2281 ELSVPARTSOIGGSSKAPSRSGSRDSTPSPAPQPLSRPIQSPGRNSISPGRNGISPPN 2340  
Qy 2341 KLSQLPRTSSPSTASTKSSGSKVSYTSPGRMQQNLTKOTGLSKNASSIPRSESASKG 2400  
Db 2341 KLSQLPRTSSPSTASTKSSGSKVSYTSPGRMQQNLTKOTGLSKNASSIPRSESASKG 2400  
Qy 2401 LQNMNNGANKKVELSRMSTKSSGSDSRPVLVROSTFIKEAPSPTLRKLLEESA 2460  
Db 2401 LQNMNNGANKKVELSRMSTKSSGSDSRPVLVROSTFIKEAPSPTLRKLLEESA 2460  
Qy 2461 SPESLSPSSRPASPTRSQATPVLSPSLPDMSLSTHSSVQAGWKLPPNLSPTEIYNDG 2520  
Db 2461 SPESLSPSSRPASPTRSQATPVLSPSLPDMSLSTHSSVQAGWKLPPNLSPTEIYNDG 2520  
Qy 2521 RPAKHDTARSHSPSRLPINRSGTWKRSHKSSSLPRVSTWRTGSSSSILSASES 2580  
Db 2521 RPAKHDTARSHSPSRLPINRSGTWKRSHKSSSLPRVSTWRTGSSSSILSASES 2580  
Qy 2581 SEKAKSEDEKHVNSISGTKQSKENQVSAKGTWRKIKENERFPTNSTQTVSSGATNGAES 2640  
Db 2581 SEKAKSEDEKHVNSISGTKQSKENQVSAKGTWRKIKENERFPTNSTQTVSSGATNGAES 2640  
Qy 2641 KTLIYQMAPAVSKTEDVWRIEDCPINNPRSGRPTGNTPPVIDSVSEKANPNIKDSKN 2700  
Db 2641 KTLIYQMAPAVSKTEDVWRIEDCPINNPRSGRPTGNTPPVIDSVSEKANPNIKDSKN 2700  
Qy 2701 QAKQVNGVSPMRTVGLNRLNSFIOVDAPDQKGTIKPGQNNPVPVSETNESSIVERT 2760  
Db 2701 QAKQVNGVSPMRTVGLNRLNSFIOVDAPDQKGTIKPGQNNPVPVSETNESSIVERT 2760  
Qy 2761 PFSSSSSKHSSPGTVAARVTPFNYPNPSPKSSADSTSARPSQIPTPVNNNTKKRDSKT 2820  
Db 2761 PFSSSSSKHSSPGTVAARVTPFNYPNPSPKSSADSTSARPSQIPTPVNNNTKKRDSKT 2820  
Qy 2821 DSTESSGTCSPKSHSGSYLVTSV 2843  
Db 2821 DSTESSGTCSPKSHSGSYLVTSV 2843

## RESULT 4

US-08-450-582-7

; Sequence 7, Application US/08450582

; Patent No. 6114124

## GENERAL INFORMATION:

; APPLICANT: ALBERTSEN, HANS

; APPLICANT: ANAND, RAKESH

; APPLICANT: CARLSON, MARY

; APPLICANT: GRODEN, JOANNA

; APPLICANT: HEDGE, PHILIP J.

; APPLICANT: JOSLYN, GEOFF

; APPLICANT: KINZLER, KENNETH

; APPLICANT: MARKHAM, ALEXANDER F.

; APPLICANT: NAKAMURA, YUSUKE

; APPLICANT: THLIVERIS, ANDREW

; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC

; NUMBER OF SEQUENCES: 102

; CORRESPONDENCE ADDRESS:

ADDRESSEE: Banner & Witcoff, Ltd.  
STREET: 1001 G Street, NW  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20001-4598  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/450,582  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/452,655  
FILING DATE: 25-MAY-1995  
APPLICATION NUMBER: US 08/289,548  
FILING DATE: 12-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/741,940  
FILING DATE: 08-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Kagan, Sarah A.  
REGISTRATION NUMBER: 32,141  
REFERENCE/DOCKET NUMBER: 1107.49964  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-508-9100  
TELEFAX: 202-508-9299  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2843 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ANTI-SENSE: NO  
US-08-450-582-7

Query Match 100.0%; Score 14566; DB 3; Length 2843;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2843; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MAAASYDQLLKQVEALKKXENSLRQELDNLSNHLTKLETSANMKVEVLKQLOGSIEDAM	60
DB	1	MAAASYDQLLKQVEALKKXENSLRQELDNLSNHLTKLETSANMKVEVLKQLOGSIEDAM	60
QY	61	ASSGQIDLLERKELNLDSSNPPGVKLRSMKLSRSGREGSVSRSGECSPVPMGSPFR	120
DB	61	ASSGQIDLLERKELNLDSSNPPGVKLRSMKLSRSGREGSVSRSGECSPVPMGSPFR	120
QY	121	RGFVNGSRESTGYLEEKERSILLADLDKEKEKDWYVAQLONLTKRIDSPLTENFSL	180
DB	121	RGFVNGSRESTGYLEEKERSILLADLDKEKEKDWYVAQLONLTKRIDSPLTENFSL	180
QY	181	QDTMTROLEYEARQIRVAMEQLGTCQDMKRAORRIARIOIEKILIRQLLOSOAT	240
DB	181	QDTMTROLEYEARQIRVAMEQLGTCQDMKRAORRIARIOIEKILIRQLLOSOAT	240
QY	241	EAEERSQKHETGSHDAERQNGQGVGEINMATSGNGQGSTTRMDHETASVLSSSSTHSA	300
DB	241	EAEERSQKHETGSHDAERQNGQGVGEINMATSGNGQGSTTRMDHETASVLSSSSTHSA	300
QY	301	PRRLTSHLGTKVEMVYSLLSMLGTHDKDDMSRTLLAMSSSQSCISMRQSGCLPILLQL	360
DB	301	PRRLTSHLGTKVEMVYSLLSMLGTHDKDDMSRTLLAMSSSQSCISMRQSGCLPILLQL	360
QY	361	HGNDKDSVLLGNRSGKEARASAAALHNIHSPQDDKRGREIRVHLLEQIRAYCETC	420
DB	361	HGNDKDSVLLGNRSGKEARASAAALHNIHSPQDDKRGREIRVHLLEQIRAYCETC	420

QY	421	WEWQEAHEPGMDQDKNMPAPVEHQICPAVCVLMKLSFDEHRHAMNELGGLOAIELLO	480
DB	421	WEWQEAHEPGMDQDKNMPAPVEHQICPAVCVLMKLSFDEHRHAMNELGGLOAIELLO	480
QY	481	VDCEMYGLTNDHYSITLRRYAGMALNLTFGDVANKATLCSMKGCMRALVAOLKSSIDL	540
DB	481	VDCEMYGLTNDHYSITLRRYAGMALNLTFGDVANKATLCSMKGCMRALVAOLKSSIDL	540
QY	541	QQVTASVLRNLNLSWRADVNSKKTILREVGSVKALMECALEVKKESTLKSLSALWLSAHT	600
DB	541	QQVTASVLRNLNLSWRADVNSKKTILREVGSVKALMECALEVKKESTLKSLSALWLSAHT	600
QY	601	ENKADICAVDGAFLVGLTLYRSQNTLAIIBSGGILRNVSLLIATNEDHQIILRENN	660
DB	601	ENKADICAVDGAFLVGLTLYRSQNTLAIIBSGGILRNVSLLIATNEDHQIILRENN	660
QY	661	CLOTLLQHLKSHSLTIVSNACGLTNLSARNPKQDALMDGAVSMLKNLHSHKHMIAM	720
DB	661	CLOTLLQHLKSHSLTIVSNACGLTNLSARNPKQDALMDGAVSMLKNLHSHKHMIAM	720
QY	721	GSAALRNLMANRPAPKYKDIANIMSPGSSIPSLHVRKQKALEABLDQHLSETFDNDLS	780
DB	721	GSAALRNLMANRPAPKYKDIANIMSPGSSIPSLHVRKQKALEABLDQHLSETFDNDLS	780
QY	781	PKASHRSKQHKQSLYCDYVFDNRHDDNRSDNFTGNMTVLSPLYNTTTLVLPSSSSSRGS	840
DB	781	PKASHRSKQHKQSLYCDYVFDNRHDDNRSDNFTGNMTVLSPLYNTTTLVLPSSSSSRGS	840
QY	841	LDSSRSEKDRSLERERIGLGNYPATENPGTSSKRGLOISTTAAQIAKWEVSVAIHTS	900
DB	841	LDSSRSEKDRSLERERIGLGNYPATENPGTSSKRGLOISTTAAQIAKWEVSVAIHTS	900
QY	901	QEDRSSGSTTELHCVTDERNALRRSAAHSTNYFTKSENRRCTCSMPYAKLEYKRSS	960
DB	901	QEDRSSGSTTELHCVTDERNALRRSAAHSTNYFTKSENRRCTCSMPYAKLEYKRSS	960
QY	961	NDLSNVSSSDGYGKRGQMKPSIESYSEDDSEKFCYGYPADLAHKKIHSANEMDNDGE	1020
DB	961	NDLSNVSSSDGYGKRGQMKPSIESYSEDDSEKFCYGYPADLAHKKIHSANEMDNDGE	1020
QY	1021	LDTPINYSKYSDQLNSGRQSPQNERWAPKHIIIEDEIKQSEQRQSRNQSTTPYVTE	1080
DB	1021	LDTPINYSKYSDQLNSGRQSPQNERWAPKHIIIEDEIKQSEQRQSRNQSTTPYVTE	1080
QY	1081	STDDXHLKFPQHFQGCQCVSPYSRGANGSETNRVSGNHGINQVQSGLCQEDDYEDDKP	1140
DB	1081	STDDXHLKFPQHFQGCQCVSPYSRGANGSETNRVSGNHGINQVQSGLCQEDDYEDDKP	1140
QY	1141	TNYSERYSEEEQHHEERPTNYSIKYNEEKRVHDQPIDYSLKYATDIPSSQKQSFSPKS	1200
DB	1141	TNYSERYSEEEQHHEERPTNYSIKYNEEKRVHDQPIDYSLKYATDIPSSQKQSFSPKS	1200
QY	1201	SSGQSSKTEHMSSESTSTPSSNAKRONOLHPSAQSRSGQPKAATCKVSSINGETIQ	1260
DB	1201	SSGQSSKTEHMSSESTSTPSSNAKRONOLHPSAQSRSGQPKAATCKVSSINGETIQ	1260
QY	1261	TYCVEDTPICFSRCSLSLSAEDIEGNCQTTQEADSAANTLQIAEIKEXIGTRSAEDPV	1320
DB	1261	TYCVEDTPICFSRCSLSLSAEDIEGNCQTTQEADSAANTLQIAEIKEXIGTRSAEDPV	1320
QY	1321	SEVPAVQCHPRTKSSRLQSSLSSESARHKAVFSSGAKSPSKSGAQTPKSPPEHYVQET	1380
DB	1321	SEVPAVQCHPRTKSSRLQSSLSSESARHKAVFSSGAKSPSKSGAQTPKSPPEHYVQET	1380
QY	1381	PLMFSTRCTSVSSLDSPESRSIAGSVQSEPCGMVSGIISPSDLPSDQGMPPSRSKTTP	1440
DB	1381	PLMFSTRCTSVSSLDSPESRSIAGSVQSEPCGMVSGIISPSDLPSDQGMPPSRSKTTP	1440
QY	1441	PPPTAQTKREVPKNKAPTAEKRESGPKQAANAAVORVLPDADTLLHFATESTPDGF	1500
DB	1441	PPPTAQTKREVPKNKAPTAEKRESGPKQAANAAVORVLPDADTLLHFATESTPDGF	1500
QY	1501	SCSSLSALSLEDFPIQKDVLRIMPVQENDNGNETSEBPKNESNENQKEAEKTIIDSE	1560



Db 1501 SCSSLSALSILDEPFIQKDVLMPPVQENDNGNETESQPKESNEQKEAKTIDSE 1560  
 QY 1561 KOLLDDDDDDIELEECIIISAMTKSRKAKPAQTASKLPPPVAKPSQLPVYKLLPS 1620  
 Db 1561 KOLLDDDDDDIELEECIIISAMTKSRKAKPAQTASKLPPPVAKPSQLPVYKLLPS 1620  
 QY 1621 QNRLOPKHVSFTPGDDMPVYVCEGTPINFSTATSLDITSPNNELAAAGVGRGAQ 1680  
 Db 1621 QNRLOPKHVSFTPGDDMPVYVCEGTPINFSTATSLDITSPNNELAAAGVGRGAQ 1680  
 QY 1681 SCEPEKDDTIPTBGRSTDEACGGKTSTVTIPELDDNKAEGDILAEICINSAMPKGSHPK 1740  
 Db 1681 SGEFEKDDTIPTBGRSTDEACGGKTSTVTIPELDDNKAEGDILAEICINSAMPKGSHPK 1740  
 QY 1741 FRVKKIMDVQQAASASSAPNKQDGGKKKPTSPVKPIPQNTETRYTRVRKNADSKNNLN 1800  
 Db 1741 FRVKKIMDVQQAASASSAPNKQDGGKKKPTSPVKPIPQNTETRYTRVRKNADSKNNLN 1800  
 QY 1801 AERFVSNDKSKONLXNSKDFNDKLPNNEDVRGSAFDPSPHHYPTIEGTPYCFSRND 1860  
 Db 1801 AERFVSNDKSKONLXNSKDFNDKLPNNEDVRGSAFDPSPHHYPTIEGTPYCFSRND 1860  
 QY 1861 SLSSLDFFDDDDVLSREKAEIRKAKENKSEAKVTSHTELTNSQOSANKTQATKQPINR 1920  
 Db 1861 SLSSLDFFDDDDVLSREKAEIRKAKENKSEAKVTSHTELTNSQOSANKTQATKQPINR 1920  
 QY 1921 GQPKPILQKSTFPQSSKIDPDRGAATDEKLFQFAIENTPVCFSHNSLSLSDIDQENN 1980  
 Db 1921 GQPKPILQKSTFPQSSKIDPDRGAATDEKLFQFAIENTPVCFSHNSLSLSDIDQENN 1980  
 QY 1981 NKNEPTIKETPPDSQEPKQASGVAPKSFHVEDTPVCFSRNSLSLSIDSEDDLLQ 2040  
 Db 1981 NKNEPTIKETPPDSQEPKQASGVAPKSFHVEDTPVCFSRNSLSLSIDSEDDLLQ 2040  
 QY 2041 ECISAMPKPKKPSRLKXGDKSPRNMGGILGEDITLDLKDQRPDSEHGLSPDSENF 2100  
 Db 2041 ECISAMPKPKKPSRLKXGDKSPRNMGGILGEDITLDLKDQRPDSEHGLSPDSENF 2100  
 QY 2101 WKAIQEGANSIVSLHQAASACLRSQAASDSISLSKSGISLGSFPFHLTPDQEKPT 2160  
 Db 2101 WKAIQEGANSIVSLHQAASACLRSQAASDSISLSKSGISLGSFPFHLTPDQEKPT 2160  
 QY 2161 SNKGPRILKXGKSTLTETKIESKIGKGVKVKSLITGKVRNSBISGQMKOPLQAN 2220  
 Db 2161 SNKGPRILKXGKSTLTETKIESKIGKGVKVKSLITGKVRNSBISGQMKOPLQAN 2220  
 QY 2221 MPSISRGRTHIIPGVNRNSSSTSPVSKGPPKTPASKSPSEGQATTSPRGAKPSVK 2280  
 Db 2221 MPSISRGRTHIIPGVNRNSSSTSPVSKGPPKTPASKSPSEGQATTSPRGAKPSVK 2280  
 QY 2281 ELSPVARTSQIGSSKAPRSRSDSTPRPAQPLSRPIQSPGNSISPGENGISPN 2340  
 Db 2281 ELSPVARTSQIGSSKAPRSRSDSTPRPAQPLSRPIQSPGNSISPGENGISPN 2340  
 QY 2341 KLSQLPRTSPSTASTKSSSGRMSYTPGRQMSQNLTKQTLGSKNASSIPRSESASG 2400  
 Db 2341 KLSQLPRTSPSTASTKSSSGRMSYTPGRQMSQNLTKQTLGSKNASSIPRSESASG 2400  
 QY 2401 LQWNTNGANKVKELSRMSTKSSGESDRSRPVLVQSTFIKAPSPFTLRKLEESA 2460  
 Db 2401 LQWNTNGANKVKELSRMSTKSSGESDRSRPVLVQSTFIKAPSPFTLRKLEESA 2460  
 QY 2461 SFESLSPSSRPASPTSAQTPVLSPLPMSLTHSSVQAGWRKLPNLSFTIYNDG 2520  
 Db 2461 SFESLSPSSRPASPTSAQTPVLSPLPMSLTHSSVQAGWRKLPNLSFTIYNDG 2520  
 QY 2521 RPAKRHDIAHSES6RLPINRSGTWKREHSHSSSLPRVSTWRTGSSSILSASS 2580  
 Db 2521 RPAKRHDIAHSES6RLPINRSGTWKREHSHSSSLPRVSTWRTGSSSILSASS 2580  
 QY 2581 SEKAKSEDEKHVNSISGTKQSKENQVSAKGTWRKIKENEFSPNTNSTQTVSSGATNGAES 2640

RESULT 5  
 US-08-821-355A-7  
 ; Sequence 7, Application US/08821355A  
 ; Patent No. 5851775  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Barker, Nick  
 ; APPLICANT: Clevers, Hans  
 ; APPLICANT: Korinek, Vladimir  
 ; APPLICANT: Morin, Patrice  
 ; APPLICANT: Kinzler, Kenneth  
 ; APPLICANT: Vogelstein, Bert  
 ; APPLICANT: Sparks, Andrew  
 ; TITLE OF INVENTION: Beta Catenin, TCF-4, and APC  
 ; TITLE OF INVENTION: Interact to Prevent Cancer  
 ; NUMBER OF SEQUENCES: 11  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Banner & Witcoff, Ltd.  
 ; STREET: 1001 G Street, N.W.  
 ; CITY: Washington  
 ; STATE: DC  
 ; COUNTRY: USA  
 ; ZIP: 20001  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FASTSEQ for Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/821,355A  
 ; FILING DATE: 20-MAR-1997  
 ; CLASSIFICATION: 514  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Kagan, Sarah A  
 ; REGISTRATION NUMBER: 32,145  
 ; REFERENCE/DOCKET NUMBER: 1107.05064  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 202-508-9100  
 ; TELEFAX: 202-508-9299  
 ; TELEX: 97430 BMB UT  
 ; INFORMATION FOR SEQ ID NO: 7:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 2973 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: No. 5851775e  
 ; US-08-821-355A-7

Query Match 100.0%; Score 14566; DB 2; Length 2973;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2843; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



QY 1 MAASDYDOLLKQVEALKMENSNLROELEDNEDNHLTKLETEASNNKEVLKOLQGSIEDEAM 60  
 Db 1 MAASDYDOLLKQVEALJOENENLROELEDNEDNHLTKLETEASNNKEVLKOLQGSIEDEAM 60  
 QY 61 ASSGOIDILLERLKEINLDSNPPGVKURSKMSLSYSGREGSVSSRSGECSPPVPMGSPFR 120  
 Db 61 ASSGOIDILLERLKEINLDSNPPGVKURSKMSLSYSGREGSVSSRSGECSPPVPMGSPFR 120  
 QY 121 RGFVNGSRESTGYLEFELEKESLILLADLKEKEKOWYIAQONLTKRIDSLPLTENFSL 180  
 Db 121 RGFVNGSRESTGYLEFELEKESLILLADLKEKEKOWYIAQONLTKRIDSLPLTENFSL 180  
 QY 181 QDTRRQLEYEARQIRVAMEEQLGTCODMEKRAQRRIARIQOIEKOLIRIQLQSOAT 240  
 Db 181 QDTRRQLEYEARQIRVAMEEQLGTCODMEKRAQRRIARIQOIEKOLIRIQLQSOAT 240  
 QY 241 EARSSONKHETGSHDAERONEGOVGEINMATSGNGOGSTTRMDHETASVLSSSSTHSA 300  
 Db 241 EARSSONKHETGSHDAERONEGOVGEINMATSGNGOGSTTRMDHETASVLSSSSTHSA 300  
 QY 301 PRRLTSHLGTKEVEMYSLMLGTHDKDDMSRTLLAMSSODSDCTSMFQSCCLPLLIQLL 360  
 Db 301 PRRLTSHLGTKEVEMYSLMLGTHDKDDMSRTLLAMSSODSDCTSMFQSCCLPLLIQLL 360  
 QY 361 HGNDKDSVLLNGSRGSKARAPASAAALHNIHSQPDYGRREIRVLHLLQIRAYCETC 420  
 Db 361 HGNDKDSVLLNGSRGSKARAPASAAALHNIHSQPDYGRREIRVLHLLQIRAYCETC 420  
 QY 421 WEQBAHBPBGMDQDNMPAPVEHOICPAVCVLMKLSFDEEHRHAMNELGLQIAELLO 480  
 Db 421 WEQBAHBPBGMDQDNMPAPVEHOICPAVCVLMKLSFDEEHRHAMNELGLQIAELLO 480  
 QY 481 VDCEMYGLTNDHYSITLRYAGMALTNLTFGDVANKATLCNSKGMALVAQLKSEBSDL 540  
 Db 481 VDCEMYGLTNDHYSITLRYAGMALTNLTFGDVANKATLCNSKGMALVAQLKSEBSDL 540  
 QY 541 QOVIASVLNLSWRADVNSKTLREVGSKALMECALEVKEESTLKSVLKALWNLASHCT 600  
 Db 541 QOVIASVLNLSWRADVNSKTLREVGSKALMECALEVKEESTLKSVLKALWNLASHCT 600  
 QY 601 ENKADICAVDQALAFVGLTLYRSQNTLAIIESGGILRVNSSLIATNEDHROILRENN 660  
 Db 601 ENKADICAVDQALAFVGLTLYRSQNTLAIIESGGILRVNSSLIATNEDHROILRENN 660  
 QY 661 CLQTLLOHLKSHSLTIVSNACGLTNLSARNPKDQALWDMGAVSMLKNLTHSKHKMTAM 720  
 Db 661 CLQTLLOHLKSHSLTIVSNACGLTNLSARNPKDQALWDMGAVSMLKNLTHSKHKMTAM 720  
 QY 721 GSAALRLNLMANRPKYKDANTMSPGSLPSLHVAKQKALAEALDAQHLSETFFDNILS 780  
 Db 721 GSAALRLNLMANRPKYKDANTMSPGSLPSLHVAKQKALAEALDAQHLSETFFDNILS 780  
 QY 781 PKASHRSKORHKQSLYGDYVFTNRRHDDNRSNFTNGMTVLSPYLNTTIVLPSSSSSRGS 840  
 Db 781 PKASHRSKORHKQSLYGDYVFTNRRHDDNRSNFTNGMTVLSPYLNTTIVLPSSSSSRGS 840  
 QY 841 LDSSRSKDRSLERBERGIGLGNYPHATENPGTSSKRGLOISTTAAQIAKWEEVSAIHTS 900  
 Db 841 LDSSRSKDRSLERBERGIGLGNYPHATENPGTSSKRGLOISTTAAQIAKWEEVSAIHTS 900  
 QY 901 QDRSSSGSTTELHCVTDERNALRRSAAHTNFTYNTFKSENSNRCTCMYPKALEYKRESS 960  
 Db 901 QDRSSSGSTTELHCVTDERNALRRSAAHTNFTYNTFKSENSNRCTCMYPKALEYKRESS 960  
 QY 961 NDSLNSVSSSDGYKRGQKQKPSIESYEDDESDFCSYGYPADLAHKIHSANHMDNDGE 1020  
 Db 961 NDSLNSVSSSDGYKRGQKQKPSIESYEDDESDFCSYGYPADLAHKIHSANHMDNDGE 1020  
 QY 1021 LUTPYNLSKYDQOLNSGRQSPQNERWAPKHIIEDEIKQSEORQSRQSNQSTTYPVYTE 1080  
 Db 1021 LUTPYNLSKYDQOLNSGRQSPQNERWAPKHIIEDEIKQSEORQSRQSNQSTTYPVYTE 1080

QY 1081 STDDKHLKFPQPHGQOECVSPYRSRGANGSETNVRVGSNHEGINQNVQSLSLCOEDDYEDDKP 1140  
 Db 1081 STDDKHLKFPQPHGQOECVSPYRSRGANGSETNVRVGSNHEGINQNVQSLSLCOEDDYEDDKP 1140  
 QY 1141 TNYSEYSEEEQHEERERTNYSIKYNEEKHRVDQIDIDYSLKYATDIPSSOKOSFSPKS 1200  
 Db 1141 TNYSEYSEEEQHEERERTNYSIKYNEEKHRVDQIDIDYSLKYATDIPSSOKOSFSPKS 1200  
 QY 1201 SSCQSKTTHMSSSSSNTSTPSSNAKRNQLHPSSAQSRSGOPQKAATCKVSSINQETIQ 1260  
 Db 1201 SSCQSKTTHMSSSSSNTSTPSSNAKRNQLHPSSAQSRSGOPQKAATCKVSSINQETIQ 1260  
 QY 1261 TYCVEDTPTCFRCSLSLSAEDBIGCQNTQOEADSNANTLOIAIEKIKIGTRSAEDPV 1320  
 Db 1261 TYCVEDTPTCFRCSLSLSAEDBIGCQNTQOEADSNANTLOIAIEKIKIGTRSAEDPV 1320  
 QY 1321 SEVPAYVQSPRTKSRLOQSSLSSESARHKAVEFFSSCAKSPSKSGAQTPKSPPEHYVQET 1380  
 Db 1321 SEVPAYVQSPRTKSRLOQSSLSSESARHKAVEFFSSCAKSPSKSGAQTPKSPPEHYVQET 1380  
 QY 1381 PLMFSRCTSVSSLDSPESRSIASSVQSEPCSGMVSGIISPSDLDPSPGQTMPPSRKSTPP 1440  
 Db 1381 PLMFSRCTSVSSLDSPESRSIASSVQSEPCSGMVSGIISPSDLDPSPGQTMPPSRKSTPP 1440  
 QY 1441 PPQTAQTKREVPKNKAAPTAEKEESGPKOAAVNAAVQRVQVLPDADTLHLHFAESTPDGF 1500  
 Db 1441 PPQTAQTKREVPKNKAAPTAEKEESGPKOAAVNAAVQRVQVLPDADTLHLHFAESTPDGF 1500  
 QY 1501 SCSSLSALSALDEPFIQKOVLEIMPPVQENDNGNETESEQPKESNENQKEAEKIDSE 1560  
 Db 1501 SCSSLSALSALDEPFIQKOVLEIMPPVQENDNGNETESEQPKESNENQKEAEKIDSE 1560  
 QY 1561 KOLLDDSDDDDIIEBECIISAMPTKSSKAKKPAOTASKLPPPVARKSOLPVYKLLPS 1620  
 Db 1561 KOLLDDSDDDDIIEBECIISAMPTKSSKAKKPAOTASKLPPPVARKSOLPVYKLLPS 1620  
 QY 1621 QNRLOPKHVSFTPGDDMPRVYCVETGPIINFSTATSLSDLTIESPPNELAAGSGVRGGAQ 1680  
 Db 1621 QNRLOPKHVSFTPGDDMPRVYCVETGPIINFSTATSLSDLTIESPPNELAAGSGVRGGAQ 1680  
 QY 1681 SGFEFEKEDTPTTGRSTDEAQQGKTSSVTPELDDNKAEBGDILACINSAMPKGSKHP 1740  
 Db 1681 SGFEFEKEDTPTTGRSTDEAQQGKTSSVTPELDDNKAEBGDILACINSAMPKGSKHP 1740  
 QY 1741 FRVKIMDQVQOASASSAPNKNQLOGKKKPTSPVKPIPONTEYTRVRKNADSKNNLN 1800  
 Db 1741 FRVKIMDQVQOASASSAPNKNQLOGKKKPTSPVKPIPONTEYTRVRKNADSKNNLN 1800  
 QY 1801 AERFSDNKDSKQNLKNSKDFNDKLPNNEDVRGSGFAPDSPHHYTPIEGTYPYCFSRND 1860  
 Db 1801 AERFSDNKDSKQNLKNSKDFNDKLPNNEDVRGSGFAPDSPHHYTPIEGTYPYCFSRND 1860  
 QY 1861 SLSSLDDEDDDDVLSREKAEELRKAENKSEAKVTSHTELTSNQOANKTOATAKOPINR 1920  
 Db 1861 SLSSLDDEDDDDVLSREKAEELRKAENKSEAKVTSHTELTSNQOANKTOATAKOPINR 1920  
 QY 1921 GQPKPILQKOSTPQSSKDIIPDRGAATDEKLQNFATIENTPVCFSHNSLSLSLSDIDQENN 1980  
 Db 1921 GQPKPILQKOSTPQSSKDIIPDRGAATDEKLQNFATIENTPVCFSHNSLSLSLSDIDQENN 1980  
 QY 1981 NKNEP1KETEPDPSQGEPSKPOASGYAPKSFVEDTTPVCFSRNSLSLSLSDIDEDLLQ 2040  
 Db 1981 NKNEP1KETEPDPSQGEPSKPOASGYAPKSFVEDTTPVCFSRNSLSLSLSDIDEDLLQ 2040  
 QY 2041 ECTISSAMPKPKKPSRLKGDNEKHSPRNMGGILGEDITLDLKDQRPDSEHGLSPDSENF 2100  
 Db 2041 ECTISSAMPKPKKPSRLKGDNEKHSPRNMGGILGEDITLDLKDQRPDSEHGLSPDSENF 2100  
 QY 2101 WKATQEGANSIVSLHAAAAACLSROASSDSLSLSKSGISLGSPFHLTPQOEKPFPT 2160  
 Db 2101 WKATQEGANSIVSLHAAAAACLSROASSDSLSLSKSGISLGSPFHLTPQOEKPFPT 2160  
 QY 2161 SNKGPRILKQEKSTLTETKIBESKGIKGGKVKYKSLITGKVSRSNSEISGQMKQKQLOAN 2220

```

Db 2161 SNKGPRILKPEKSTLTETKIESKGIKGGKVKYSLITGKVRNSNISEISGMKQLOAN 2220
Qy 2221 MPSTSRGRTHIHPGVNRSSTSPVSKGPPPLKTPASKSPSGQTATTSFPRGAKSVKS 2280
Db 2221 MPSTSRGRTHIHPGVNRSSTSPVSKGPPPLKTPASKSPSGQTATTSFPRGAKSVKS 2280
Qy 2281 ELSVPAQTSQIGGSKAPSRGSRDTPGRPAQPLSRPIQSPGRNISPGRNGISPPN 2340
Db 2281 ELSVPAQTSQIGGSKAPSRGSRDTPGRPAQPLSRPIQSPGRNISPGRNGISPPN 2340
Qy 2341 KLSQLPRTSPSTASTKSGSGKMSYTPGRQMSQNLTKQTGLSKNASSIPRSESASKG 2400
Db 2341 KLSQLPRTSPSTASTKSGSGKMSYTPGRQMSQNLTKQTGLSKNASSIPRSESASKG 2400
Qy 2401 LQNMNGNGANKKVELRMSSTKSGSESPPRPPVLRQSTFIKAPSPTLRRKLEESA 2460
Db 2401 LQNMNGNGANKKVELRMSSTKSGSESPPRPPVLRQSTFIKAPSPTLRRKLEESA 2460
Qy 2461 SFESLSPSRPASPTRSQATPVLSPDPMSLTHSSVQAGGWRKLPNLSPTIEYNDG 2520
Db 2461 SFESLSPSRPASPTRSQATPVLSPDPMSLTHSSVQAGGWRKLPNLSPTIEYNDG 2520
Qy 2521 RPAKRHDIAHSHSPRLPINSRGTWKRHSKSSSLPRVSTWRTTGGSSSILSASSES 2580
Db 2521 RPAKRHDIAHSHSPRLPINSRGTWKRHSKSSSLPRVSTWRTTGGSSSILSASSES 2580
Qy 2581 SERAKSEDEKHVNSISGTSKQENQVSAKGTWKIKENEPSPNTNSTQTVSSGATNGAS 2640
Db 2581 SERAKSEDEKHVNSISGTSKQENQVSAKGTWKIKENEPSPNTNSTQTVSSGATNGAS 2640
Qy 2641 KTLIYQAPAVKTEDVWVIEPCPINPRSGSPGTGNTPPVIVDSVSEKANPIKDSKON 2700
Db 2641 KTLIYQAPAVKTEDVWVIEPCPINPRSGSPGTGNTPPVIVDSVSEKANPIKDSKON 2700
Qy 2701 QAKQNVGNGVPMRTVGLNRLNSFIOVDAPDQGTETKPGQNPVPVSETNESSIVERT 2760
Db 2701 QAKQNVGNGVPMRTVGLNRLNSFIOVDAPDQGTETKPGQNPVPVSETNESSIVERT 2760
Qy 2761 PFSSSSSKHSPSGTVAARVTPFNPNPSPRKSSADSTGARPQIPTPVNNNTKRDST 2820
Db 2761 PFSSSSSKHSPSGTVAARVTPFNPNPSPRKSSADSTGARPQIPTPVNNNTKRDST 2820
Qy 2821 DSTESSGTQSPKRHSGLVTSV 2843
Db 2821 DSTESSGTQSPKRHSGLVTSV 2843

```

RESULT 6  
US-09-003-687A-7

; Sequence 7, Application US/09003687A  
; Patent No. 5998600

; GENERAL INFORMATION:

; APPLICANT: Barker, Nick

; APPLICANT: Clevers, Hans

; APPLICANT: Korinek, Vladimir

; APPLICANT: Morin, Patrice

; APPLICANT: Kinzler, Kenneth

; APPLICANT: Vogelstein, Bert

; APPLICANT: Sparks, Andrew

; TITLE OF INVENTION: Beta Catenin, TCF-4, and APC

; INTERACT TO PREVENT CANCER

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Banner & Witcoff, Ltd.

; STREET: 1001 G Street, N.W.

; CITY: Washington

; STATE: DC

; COUNTRY: USA

; ZIP: 20001

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

```

; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/003,687A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/821,355
; FILING DATE: 20-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Kagan, Sarah A
; REGISTRATION NUMBER: 32,145
; REFERENCE/DOCKET NUMBER: 1107.05064
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9299
; TELEX: 97430 EMB UT
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2973 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 5998600e
; US-09-003-687A-7

```

Query Match 100.0%; Score 14566; DB 2; Length 2973;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2843; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 MAAASYDQLLKQVQALXKXNSNLRQLEDNSNHLTKLETSANMKVYLKQLQGSIEDEAM 60
Db 1 MAAASYDQLLKQVQALXKXNSNLRQLEDNSNHLTKLETSANMKVYLKQLQGSIEDEAM 60
Qy 61 ASSGQIDLLERLKLNDSSNFPQVKLRKMSLRSYSGREGSVSSRSGEGSPVPMGSPFR 120
Db 61 ASSGQIDLLERLKLNDSSNFPQVKLRKMSLRSYSGREGSVSSRSGEGSPVPMGSPFR 120
Qy 121 RGFVNGRESTGYLEELKERSLLADLDKEEKEKWYQALQNLTKRIDSLPTENFSL 180
Db 121 RGFVNGRESTGYLEELKERSLLADLDKEEKEKWYQALQNLTKRIDSLPTENFSL 180
Qy 181 QDTRRQLEYEARQIRVAMEEQLGTCQDMKRAQRRIARIQIEKDILRIQLQSOAT 240
Db 181 QDTRRQLEYEARQIRVAMEEQLGTCQDMKRAQRRIARIQIEKDILRIQLQSOAT 240
Qy 241 EAERSSQNKHETGSHDAERQNEGQVGEINNATSGNQGQSTTRMDHETASVLSSTHSA 300
Db 241 EAERSSQNKHETGSHDAERQNEGQVGEINNATSGNQGQSTTRMDHETASVLSSTHSA 300
Qy 301 PRLTSHLGTQVEMVYSLIIMLGTHDKDMSRTLLAMSSQDSCI SMRQSCCLPLILQLL 360
Db 301 PRLTSHLGTQVEMVYSLIIMLGTHDKDMSRTLLAMSSQDSCI SMRQSCCLPLILQLL 360
Qy 361 HGNDKDSVLLGNRSGSKEARASAAALHNIIHSQDDPKRGRREIRVLHLEEQIRAYCETC 420
Db 361 HGNDKDSVLLGNRSGSKEARASAAALHNIIHSQDDPKRGRREIRVLHLEEQIRAYCETC 420
Qy 421 WEQEAHPGVDQDKNPMPAPVEHQICPAVCVLMKLSFDEEHRHAMNELGGLQAIALLQ 480
Db 421 WEQEAHPGVDQDKNPMPAPVEHQICPAVCVLMKLSFDEEHRHAMNELGGLQAIALLQ 480
Qy 481 VDCEMYGLTNDHYSITLRRVAGMALTNLTGPDVANKATLCSMKGCNRAALVAQLKSESDL 540
Db 481 VDCEMYGLTNDHYSITLRRVAGMALTNLTGPDVANKATLCSMKGCNRAALVAQLKSESDL 540
Qy 541 QQVIASVLRNLSWRADVNSKKTIREVGSVKALMECALEVKKESTLKSVALNLSAHCT 600
Db 541 QQVIASVLRNLSWRADVNSKKTIREVGSVKALMECALEVKKESTLKSVALNLSAHCT 600
Qy 601 ENKADICAVDGAFLVGTITTYRSQNTLAIIESGGILRNVSLLIATNEDHQLRENN 660
Db 601 ENKADICAVDGAFLVGTITTYRSQNTLAIIESGGILRNVSLLIATNEDHQLRENN 660

```

QY 661 CLQTLLOHLKSHSLTIVSNACGTLWNLSARNPKDQALWDMGAVSMLKNLILHSHKMIAM 720  
DB 661 CLQTLLOHLKSHSLTIVSNACGTLWNLSARNPKDQALWDMGAVSMLKNLILHSHKMIAM 720  
QY 721 GSAALRNLMANRPKAKYKXANITMSPGSSLPJLHVWKQXALFAELDAQHLSTFFDNILS 780  
DB 721 GSAALRNLMANRPKAKYKXANITMSPGSSLPJLHVWKQXALFAELDAQHLSTFFDNILS 780  
QY 781 PKASHRSKQKHQSLGYDVFDTNHDNRSDNENTGNMTVLSPYLNTTTLVPSGSSSRGS 840  
DB 781 PKASHRSKQKHQSLGYDVFDTNHDNRSDNENTGNMTVLSPYLNTTTLVPSGSSSRGS 840  
QY 841 LPSRSEKDRSLERERIGLGNYPHYPATENPGTSSKRGLOISTAAQIAKWEEVSIAHTS 900  
DB 841 LPSRSEKDRSLERERIGLGNYPHYPATENPGTSSKRGLOISTAAQIAKWEEVSIAHTS 900  
QY 901 QEDRSSTTEHLCHVTDERNALRRSAATHENTYNTKSENSNETCSMPYAKLEYKRSS 960  
DB 901 QEDRSSTTEHLCHVTDERNALRRSAATHENTYNTKSENSNETCSMPYAKLEYKRSS 960  
QY 961 NDSLNSVSSDGYKRGQMKPSESYSDDESKFCYGOYPADLAHKHSANHMDNDNDE 1020  
DB 961 NDSLNSVSSDGYKRGQMKPSESYSDDESKFCYGOYPADLAHKHSANHMDNDNDE 1020  
QY 1021 LPTPINSLKYDEQLNSGRQSPQNERWARPKHIEDEIKQSEORQSRNOSTTYPVYTE 1080  
DB 1021 LPTPINSLKYDEQLNSGRQSPQNERWARPKHIEDEIKQSEORQSRNOSTTYPVYTE 1080  
QY 1081 STDDKHLKFPQHFQOQECVSPVRSRGANGSENRVGNHGINQVNSQSLCOEDDYDDKP 1140  
DB 1081 STDDKHLKFPQHFQOQECVSPVRSRGANGSENRVGNHGINQVNSQSLCOEDDYDDKP 1140  
QY 1141 TNYSRYSBEEQHEEERPTNYSIKYNEBKRVDPIDYSLKYATDIPSSQKQSFSPKS 1200  
DB 1141 TNYSRYSBEEQHEEERPTNYSIKYNEBKRVDPIDYSLKYATDIPSSQKQSFSPKS 1200  
QY 1201 SSGQSKTEHMSSENTSTPSSNAGQNQLHPSSAQSRSGOPQKAAATCKVSSINQETIQ 1260  
DB 1201 SSGQSKTEHMSSENTSTPSSNAGQNQLHPSSAQSRSGOPQKAAATCKVSSINQETIQ 1260  
QY 1261 TYCVSDTPTCFGRCSLSSLSAEDEIGCNQTOQADSAANTQIAEIKEIKIGTRSAEDPV 1320  
DB 1261 TYCVSDTPTCFGRCSLSSLSAEDEIGCNQTOQADSAANTQIAEIKEIKIGTRSAEDPV 1320  
QY 1321 SEVPASVQHPRTKSRLOQSSLSSESARHKAVEFSSGAKSPSKSGAOTPKPPEHYOET 1380  
DB 1321 SEVPASVQHPRTKSRLOQSSLSSESARHKAVEFSSGAKSPSKSGAOTPKPPEHYOET 1380  
QY 1381 PLMFRCCTSVSLDGFESRSIASSVQSEPCSGWVGIIISPSDLPSDQOTMPPSRSKTTP 1440  
DB 1381 PLMFRCCTSVSLDGFESRSIASSVQSEPCSGWVGIIISPSDLPSDQOTMPPSRSKTTP 1440  
QY 1441 PPPQTAQTKREVPKNAKPTAEKRESGPKQAAVNAAVQRVQVLPDADTLHLFATESTPDGF 1500  
DB 1441 PPPQTAQTKREVPKNAKPTAEKRESGPKQAAVNAAVQRVQVLPDADTLHLFATESTPDGF 1500  
QY 1501 SCSSSLSALSLEPFIQKDLVRLIMPVQENDNGNETESEOPKESNENOEKAEKTIIDSE 1560  
DB 1501 SCSSSLSALSLEPFIQKDLVRLIMPVQENDNGNETESEOPKESNENOEKAEKTIIDSE 1560  
QY 1561 KOLLDDSDDDDDIEILEECIIISAMPTKSRKAKKPAQTASKLPPPVARPKPSQLPVYKLLPS 1620  
DB 1561 KOLLDDSDDDDDIEILEECIIISAMPTKSRKAKKPAQTASKLPPPVARPKPSQLPVYKLLPS 1620  
QY 1621 QNRLOPKHVSFTPGDDMPRVYCVETGPIINFSTATSLDLTIESPNELAAEGVRRGAQ 1680  
DB 1621 QNRLOPKHVSFTPGDDMPRVYCVETGPIINFSTATSLDLTIESPNELAAEGVRRGAQ 1680  
QY 1681 SGEFEKRDITPTEGRSTDEAQGGKTSVTPILDDNKAEGDILAEICINSAMPKGSHPK 1740  
DB 1681 SGEFEKRDITPTEGRSTDEAQGGKTSVTPILDDNKAEGDILAEICINSAMPKGSHPK 1740

QY 1741 FRVKKIMDOVQOASASSAPNKNQLOGKKKPTSPVKPIPQNTFYTRVKNADSKNNLN 1800  
DB 1741 FRVKKIMDOVQOASASSAPNKNQLOGKKKPTSPVKPIPQNTFYTRVKNADSKNNLN 1800  
QY 1801 AERVFSNKNDSKQNLKNNSKDFNDKLPNNEDVRGSPAFDPSHHYTPISGTYCFSRND 1860  
DB 1801 AERVFSNKNDSKQNLKNNSKDFNDKLPNNEDVRGSPAFDPSHHYTPISGTYCFSRND 1860  
QY 1861 SLGSLDEDDDDVLSRKAELRKAENKESAEKVTSHTELTSNQQANKTQALAKQPINR 1920  
DB 1861 SLGSLDEDDDDVLSRKAELRKAENKESAEKVTSHTELTSNQQANKTQALAKQPINR 1920  
QY 1921 GQPKPILQKQSTFPQSSKDIIPDRGAATDEKLQNAFAIENTPVCFSHNSLSLSDIDQENN 1980  
DB 1921 GQPKPILQKQSTFPQSSKDIIPDRGAATDEKLQNAFAIENTPVCFSHNSLSLSDIDQENN 1980  
QY 1981 NKENEPKETEPDPDSQGEPSKPOASGYAPKSFHVEDTTPVCFSRNSLSLSLUSIUSEDLLQ 2040  
DB 1981 NKENEPKETEPDPDSQGEPSKPOASGYAPKSFHVEDTTPVCFSRNSLSLSLUSIUSEDLLQ 2040  
QY 2041 ECISSAMPKPKKPSRLKGDNEKESPRNMGGILGEDITLCLKDQRPDSEHGLSPDSENF 2100  
DB 2041 ECISSAMPKPKKPSRLKGDNEKESPRNMGGILGEDITLCLKDQRPDSEHGLSPDSENF 2100  
QY 2101 WKAIQEGANSIVSSLHQAACCLSRQASDSDSIIISLKSGLISGSPPHLTPDQEEKPFT 2160  
DB 2101 WKAIQEGANSIVSSLHQAACCLSRQASDSDSIIISLKSGLISGSPPHLTPDQEEKPFT 2160  
QY 2161 SNKGPRILKPEGEKSTLTETKIESESKGKGGKVKYKSLITGKVRNSSEISGQNKQPOLOAN 2220  
DB 2161 SNKGPRILKPEGEKSTLTETKIESESKGKGGKVKYKSLITGKVRNSSEISGQNKQPOLOAN 2220  
QY 2221 MPSISRORTMIHIPGVNNSSTSPVSKGPPKPTASKPSPEGOTATTSPRAKPSVKS 2280  
DB 2221 MPSISRORTMIHIPGVNNSSTSPVSKGPPKPTASKPSPEGOTATTSPRAKPSVKS 2280  
QY 2281 ELSVARSQTSQIGGSSKAPSRSGSRDTPSRPAQOPLSRPIQSPGRNSISPGRANGISPPN 2340  
DB 2281 ELSVARSQTSQIGGSSKAPSRSGSRDTPSRPAQOPLSRPIQSPGRNSISPGRANGISPPN 2340  
QY 2341 KLSQLPRTSPSTASTKSSGSGKMSYTPSGRQMSQQNLTKQTLGKSNASSIPRSESASKG 2400  
DB 2341 KLSQLPRTSPSTASTKSSGSGKMSYTPSGRQMSQQNLTKQTLGKSNASSIPRSESASKG 2400  
QY 2401 LNQMANGANKKVELSEMSTKSSGSESDSRERPVLRQSTTIKEAPSPTLARKLEESA 2460  
DB 2401 LNQMANGANKKVELSEMSTKSSGSESDSRERPVLRQSTTIKEAPSPTLARKLEESA 2460  
QY 2461 SFESLSPSRPASPTRSQAQTPVLSPSLPDMSLSTHSSVQAGGWRKLPPLNLSPTIEYNDG 2520  
DB 2461 SFESLSPSRPASPTRSQAQTPVLSPSLPDMSLSTHSSVQAGGWRKLPPLNLSPTIEYNDG 2520  
QY 2521 RPAKRHDIAKSHSESPRLPINRSGTWKREHSHKSSSLPRVSTWRRRTGSSSSILSASSES 2580  
DB 2521 RPAKRHDIAKSHSESPRLPINRSGTWKREHSHKSSSLPRVSTWRRRTGSSSSILSASSES 2580  
QY 2581 SEKAKSDEKHNVSISGTKOSKENOVSAKGTWRKIKENEFSPNTSQTSSGATNGAES 2640  
DB 2581 SEKAKSDEKHNVSISGTKOSKENOVSAKGTWRKIKENEFSPNTSQTSSGATNGAES 2640  
QY 2641 KTLIYQMAPAVSKTDEYVWRIEDCPIINNRSGSPGNTGTPVIDSVSEKANPNIKOSKDN 2700  
DB 2641 KTLIYQMAPAVSKTDEYVWRIEDCPIINNRSGSPGNTGTPVIDSVSEKANPNIKOSKDN 2700  
QY 2701 QAKQNVGSGVPWRTVGLNRLNSFIQVADPDQGTETIKQGNPNVPVSTNESSIVERT 2760  
DB 2701 QAKQNVGSGVPWRTVGLNRLNSFIQVADPDQGTETIKQGNPNVPVSTNESSIVERT 2760  
QY 2761 PFSSSSSSSKHSPSGTVAARVTFPNYNPSPRKSSADSTSARPSQIPTPVNNNTKKRDSKT 2820  
DB 2761 PFSSSSSSSKHSPSGTVAARVTFPNYNPSPRKSSADSTSARPSQIPTPVNNNTKKRDSKT 2820  
QY 2821 DSTESSGTQSPKRRHSGSYLVTSV 2843



Db 2761 PFSSSSSKHSPSGTVAARVTPFNPNPGRKSSADTSARPSQIPTPVNNNTKRDST 2820

Qy 2821 DSTESSGTQSPKRHSGLYLTSV 2843

Db 2821 DSTESSGTQSPKRHSGLYLTSV 2843

RESULT 8

US-07-741-940-7

; Sequence 7, Application US/07741940

; Patent No. 5352775

; GENERAL INFORMATION:

; APPLICANT: ALBERTSEN, HANS

; APPLICANT: ANAND, RAKESH

; APPLICANT: CARLSON, MARY

; APPLICANT: GRODEN, JOANNA

; APPLICANT: HEDGE, PHILIP J.

; APPLICANT: JOSLYN, GEOFF

; APPLICANT: KINZLER, KENNETH

; APPLICANT: MARKHAM, ALEXANDER F.

; APPLICANT: NAKAMURA, YUSUKE

; APPLICANT: THILVERIS, ANDREW

; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC

; TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS

; NUMBER OF SEQUENCES: 94

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Banner, Birch, McKie & Beckett

; STREET: 1001 G Street, NW

; CITY: Washington

; STATE: D.C.

; COUNTRY: USA

; ZIP: 20001-4598

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/741,940

; FILING DATE: 19920109

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Kagan, Sarah A.

; REGISTRATION NUMBER: 32,141

; REFERENCE/DOCKET NUMBER: 1107.035574

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 202-508-9100

; TELEFAX: 202-508-9299

; INFORMATION FOR SEQ ID NO: 7:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2842 amino acids

; TYPE: AMINO ACID

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; ORIGINAL SOURCE:

; ORGANISM: Homo sapiens

; IMMEDIATE SOURCE:

; CLONE: APC

US-07-741-940-7

Query Match 99.9%; Score 14548.5; DB 1; Length 2842;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2842; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 MAASVDLLKQVEALKMENSRLRQLEDNSNHLTKLETEASNKVELKQLOGSIDEAM 60

Db 1 MAASVDLLKQVEALKMENSRLRQLEDNSNHLTKLETEASNKVELKQLOGSIDEAM 60

Qy 61 ASSQIDLLRLKXELNLDSSNFFGVKLRKQWLSRYSRSGSVSSRSGECSVPVPMGSFPR 120

Db 61 ASSQIDLLRLKXELNLDSSNFFGVKLRKQWLSRYSRSGSVSSRSGECSVPVPMGSFPR 120

Db 1681 SGFEFKRDTPTPEGRSTDEAQGKTSVVTIPELDDNKAEEDGILAEICINSAMPKSHKP 1740

Qy 1741 FRVKIMQVQOQASASSAPKNQOLDGKKKPTSPVKPIPONTEYRTVRKNADSKNLN 1800

Db 1741 FRVKIMQVQOQASASSAPKNQOLDGKKKPTSPVKPIPONTEYRTVRKNADSKNLN 1800

Qy 1801 AERFSDNKKSKQNLKNNKDFNDKLPNNEDRVGSAFDPSPHHYPIEGTPYCFSEND 1860

Db 1801 AERFSDNKKSKQNLKNNKDFNDKLPNNEDRVGSAFDPSPHHYPIEGTPYCFSEND 1860

Qy 1861 SLSSLDPDODDVLDSREKAEURLKAKENKESAKYTSHTELTSNQSANKTQAIKQPINR 1920

Db 1861 SLSSLDPDODDVLDSREKAEURLKAKENKESAKYTSHTELTSNQSANKTQAIKQPINR 1920

Qy 1921 GQPKILQKSTFFOSSKDIPDRGAATDEKLNFAIENTPVCFSHNSLSLSISDIQDQNN 1980

Db 1921 GQPKILQKSTFFOSSKDIPDRGAATDEKLNFAIENTPVCFSHNSLSLSISDIQDQNN 1980

Qy 1981 NKENPIKETEPDPSQGPSPKPAQAGYAPKSPHVEDTTPVCFSRNSLSLSISDEDDLLQ 2040

Db 1981 NKENPIKETEPDPSQGPSPKPAQAGYAPKSPHVEDTTPVCFSRNSLSLSISDEDDLLQ 2040

Qy 2041 ECISAMPKKKPSRLKGNENKHSRNMGGILGEDLTLDLKDIOQPDSEHGLSPDSENF 2100

Db 2041 ECISAMPKKKPSRLKGNENKHSRNMGGILGEDLTLDLKDIOQPDSEHGLSPDSENF 2100

Qy 2101 WKAIQEGANSIVSSLHQAAAAACLSRQASPSDSILSLKSGISLSPHLPDQEKPT 2160

Db 2101 WKAIQEGANSIVSSLHQAAAAACLSRQASPSDSILSLKSGISLSPHLPDQEKPT 2160

Qy 2161 SNKGRILKPKGKSTLETWKLI3ESGKIGKGGKVYKSLITCKVRNSLSISGMQKPLQAN 2220

Db 2161 SNKGRILKPKGKSTLETWKLI3ESGKIGKGGKVYKSLITCKVRNSLSISGMQKPLQAN 2220

Qy 2221 MFSISGRRTMIHPGVNRSSSTSPVSKKGPPLKTPASKSPSEGTATTSRPGAKPSVKS 2280

Db 2221 MFSISGRRTMIHPGVNRSSSTSPVSKKGPPLKTPASKSPSEGTATTSRPGAKPSVKS 2280

Qy 2281 ELSPVARTQTSIGGSKAPSRSGSDSTPSRPAQPLSRPIQSPGRNSISFGRNGISPPN 2340

Db 2281 ELSPVARTQTSIGGSKAPSRSGSDSTPSRPAQPLSRPIQSPGRNSISFGRNGISPPN 2340

Qy 2341 KLSQLPRTSSPTAGTKSGSGKMYTSPGRMSQONLTQTGLSKNASSIPRESASKG 2400

Db 2341 KLSQLPRTSSPTAGTKSGSGKMYTSPGRMSQONLTQTGLSKNASSIPRESASKG 2400

Qy 2401 LNMNNGANKKVELSRMSTKSGSGSDRSERPVLRQSTTFKEAPSPTLRRKLEESA 2460

Db 2401 LNMNNGANKKVELSRMSTKSGSGSDRSERPVLRQSTTFKEAPSPTLRRKLEESA 2460

Qy 2461 SFESLSPSRPASPTRSOAQTPLSPSLPDYSLTHSSVQAGGWRKLPNLSPTIEYNDG 2520

Db 2461 SFESLSPSRPASPTRSOAQTPLSPSLPDYSLTHSSVQAGGWRKLPNLSPTIEYNDG 2520

Qy 2521 RPAKRHDIAHSHSESPRLPINRSGTWKREHSHKSSSLPRVSTWRRITGSSSSILSASSES 2580

Db 2521 RPAKRHDIAHSHSESPRLPINRSGTWKREHSHKSSSLPRVSTWRRITGSSSSILSASSES 2580

Qy 2581 SEKAKSEBKHVNSISGTKQSKENQVSAKGTWRKIKENEFPTNSTQTSVSGATNGAES 2640

Db 2581 SEKAKSEBKHVNSISGTKQSKENQVSAKGTWRKIKENEFPTNSTQTSVSGATNGAES 2640

Qy 2641 KTLIYQMAPAVAKTIDVWYRIEDCPINNPRSGRSPGTNTPPVIDSVSEKAPNIKDKDN 2700

Db 2641 KTLIYQMAPAVAKTIDVWYRIEDCPINNPRSGRSPGTNTPPVIDSVSEKAPNIKDKDN 2700

Qy 2701 QAKQNVGSGVPMRTVGLNRLNLSFTQVDAPQKGTETIKPGQNNPVPVSETNESSIVERT 2760

Db 2701 QAKQNVGSGVPMRTVGLNRLNLSFTQVDAPQKGTETIKPGQNNPVPVSETNESSIVERT 2760

Qy 2761 PFSSSSSKHSPSGTVAARVTPFNPNPGRKSSADTSARPSQIPTPVNNNTKRDST 2820

Db 2761 PFSSSSSKHSPSGTVAARVTPFNPNPGRKSSADTSARPSQIPTPVNNNTKRDST 2820

121 RGVNGSRSTGYLEELKERSLILLADLKEEKEDWYVAQLONLTKRIDSLPLTENFSL 180  
 121 RGVNGSRSTGYLEELKERSLILLADLKEEKEDWYVAQLONLTKRIDSL- L1TENFSL 179  
 181 QTMTRQLBYEARQIRVAMEBOLGTCQDMKRAQRIARIQOIEKDILRIROLLOSQAT 240  
 180 QTMTRQLBYEARQIRVAMEBOLGTCQDMKRAQRIARIQOIEKDILRIROLLOSQAT 239  
 241 EARSSONKHETGSHDAERQEGGVEINMATSGNGQSGSTTRMDHETASVLSSTSTHGA 300  
 240 EARSSONKHETGSHDAERQEGGVEINMATSGNGQSGSTTRMDHETASVLSSTSTHGA 299  
 301 PRRLTSHLGTKEVYVSLLSMLGTHDKDMSRRTLLAMSSQDSCINRQSGCLPLLIQLL 360  
 300 PRRLTSHLGTKEVYVSLLSMLGTHDKDMSRRTLLAMSSQDSCINRQSGCLPLLIQLL 359  
 361 HGNDKOSVLGNRSGSKEARASAAALHNIIHSQDDKGRREIRVHLLEQIRAYCETC 420  
 360 HGNDKOSVLGNRSGSKEARASAAALHNIIHSQDDKGRREIRVHLLEQIRAYCETC 419  
 421 WEMQEAHEPGWQDKNPMPAPVEHQICPAVCVLMKLSFDEEHRHAMNELGGLQAI AELLQ 480  
 420 WEMQEAHEPGWQDKNPMPAPVEHQICPAVCVLMKLSFDEEHRHAMNELGGLQAI AELLQ 479  
 481 VDCMYGLTNDHYSITLRRYAGHALTNLTGPDVANKATLCSMKGCMRALVAQLKSEBDL 540  
 480 VDCMYGLTNDHYSITLRRYAGHALTNLTGPDVANKATLCSMKGCMRALVAQLKSEBDL 539  
 541 QOVLASVLRNLWRADVNSKKTREVGSVKALMECALEVKKESTLXSVLSALWNLSAHCT 600  
 540 QOVLASVLRNLWRADVNSKKTREVGSVKALMECALEVKKESTLXSVLSALWNLSAHCT 599  
 601 ENKADI CAVDGALAFVLGTYRSQTNLTALIESGGGILNVSSLIATNEDHRIQLRENN 660  
 600 ENKADI CAVDGALAFVLGTYRSQTNLTALIESGGGILNVSSLIATNEDHRIQLRENN 659  
 661 CLOTLLOHLKSHSLTIVSNACGLTNLWLSARNPKDQALWDMGAVSMKMLNLIHSHKHMIAM 720  
 660 CLOTLLOHLKSHSLTIVSNACGLTNLWLSARNPKDQALWDMGAVSMKMLNLIHSHKHMIAM 719  
 721 GSAALRNLMANRPAXYKDIANIMSPGSLPSLHVRKQKALEAELDAQHLSSETFDNIDNLS 780  
 720 GSAALRNLMANRPAXYKDIANIMSPGSLPSLHVRKQKALEAELDAQHLSSETFDNIDNLS 779  
 781 PKASHRSKQBKOSLYGQVDFDNRHDDNBSDFNCTGNMTVLSPLYNTTVLPSSSSSRGS 840  
 780 PKASHRSKQBKOSLYGQVDFDNRHDDNBSDFNCTGNMTVLSPLYNTTVLPSSSSSRGS 839  
 841 LDSRSSEKSLREBERGIGLGNVHPATENPGTSSKRGLOISTTAAQIAKVMEEVSAIHTS 900  
 840 LDSRSSEKSLREBERGIGLGNVHPATENPGTSSKRGLOISTTAAQIAKVMEEVSAIHTS 899  
 901 QEDRSSGTTLHCVTDERNALRSSAAHSHNTYFTKSENSNRCTCSMPYAKLEYKRS 960  
 900 QEDRSSGTTLHCVTDERNALRSSAAHSHNTYFTKSENSNRCTCSMPYAKLEYKRS 959  
 961 NDSLNSVSSDGYGKQOMKPSIESYEDDESFCVSGQVPADLAHKIHSANHMDNDGE 1020  
 960 NDSLNSVSSDGYGKQOMKPSIESYEDDESFCVSGQVPADLAHKIHSANHMDNDGE 1019  
 1021 LDTPINYSKYSDQLNGRQSPQSNRWARPKHIIIEDEIKQSEQRNQSTTYPVYTE 1080  
 1020 LDTPINYSKYSDQLNGRQSPQSNRWARPKHIIIEDEIKQSEQRNQSTTYPVYTE 1079  
 1081 STDDKHLKQFQHQOCVSPYRSRGANGSETNRVGNHGINOVQSLOEODDYEDDKP 1140  
 1080 STDDKHLKQFQHQOCVSPYRSRGANGSETNRVGNHGINOVQSLOEODDYEDDKP 1139  
 1141 TNSYERYSEEQHEEERPTNYSIKYNEEKRVDPIDYSLKYATDIPSSQKQSFSSKS 1200  
 1140 TNSYERYSEEQHEEERPTNYSIKYNEEKRVDPIDYSLKYATDIPSSQKQSFSSKS 1199  
 1201 SSCOSSYTHWSSSPTWSTPSSNAKZONOLHPS3QSP5QCBKAAATCKVSSINQETIQ 1260

1200 SSGQSSKTEHSSSSSENTSTPSSNAKZONQLHPS3QSP5QCBKAAATCKVSSINQETIQ 1259  
 1261 TYCYVEDTPICFSRCSLSLSLSAEDGICNOTTOEADSANTLQIAEIKKIGTRSAADPV 1320  
 1260 TYCYVEDTPICFSRCSLSLSLSAEDGICNOTTOEADSANTLQIAEIKKIGTRSAADPV 1319  
 1321 SEVPAVSQHPRTKSRRLQSSLSSESARHKAVERFSSGAKSPSKSGAQTTPSPPHYVOET 1380  
 1320 SEVPAVSQHPRTKSRRLQSSLSSESARHKAVERFSSGAKSPSKSGAQTTPSPPHYVOET 1379  
 1381 PLMFSRCTSVSSLOSFSRSSTASSVQSEPCSGMVGIIISPSDLPDSPGQTMPPSRSTPP 1440  
 1380 PLMFSRCTSVSSLOSFSRSSTASSVQSEPCSGMVGIIISPSDLPDSPGQTMPPSRSTPP 1439  
 1441 PPPOTACTKREVPKXKAPTAEKRESGPKQAANVAQVQVLPDADTLLHFATESTPDGF 1500  
 1440 PPPOTACTKREVPKXKAPTAEKRESGPKQAANVAQVQVLPDADTLLHFATESTPDGF 1499  
 1501 SCSSLSLSLDEPFIQKVELRIMPPVOENDNGNETESOPKESNENQKBAKTTIDSE 1560  
 1500 SCSSLSLSLDEPFIQKVELRIMPPVOENDNGNETESOPKESNENQKBAKTTIDSE 1559  
 1561 KOLLDDSDDDIIELEECIISAMPTKSRKAKKPAQATASKLPPPVARKPSOLPYKLLPS 1620  
 1560 KOLLDDSDDDIIELEECIISAMPTKSRKAKKPAQATASKLPPPVARKPSOLPYKLLPS 1619  
 1621 QNRLOPKHVSFTPGDDMPRVYCVEGTPIINPSTATSLSDLTIESPPNELAAGEVGRGAQ 1680  
 1620 QNRLOPKHVSFTPGDDMPRVYCVEGTPIINPSTATSLSDLTIESPPNELAAGEVGRGAQ 1679  
 1681 SGEFEKRTITIEGSRSTDEAOGGKTSVTIPELDDNKAEGDILAEICINSAMPKKGSHKP 1740  
 1680 SGEFEKRTITIEGSRSTDEAOGGKTSVTIPELDDNKAEGDILAEICINSAMPKKGSHKP 1739  
 1741 FRVKKIMDQVQOASASSAPKNQLDGKKKPTSPVKPIPONTFYRTRVRKNADSKNNLN 1800  
 1740 FRVKKIMDQVQOASASSAPKNQLDGKKKPTSPVKPIPONTFYRTRVRKNADSKNNLN 1799  
 1801 AERVFSDNKDOKKQNLKNNKDFNDKLPNNEDRVGSGFADSPHHYTPIEGTPYCFSRND 1860  
 1800 AERVFSDNKDOKKQNLKNNKDFNDKLPNNEDRVGSGFADSPHHYTPIEGTPYCFSRND 1859  
 1861 SLSLDFDQDDVDLSREKAELRKAKENKESAKVTSHTELTSNQQSANKTQAIJAKQPINR 1920  
 1860 SLSLDFDQDDVDLSREKAELRKAKENKESAKVTSHTELTSNQQSANKTQAIJAKQPINR 1919  
 1921 GQPKILOKQSTFPQSSKDIPDRGAATDEKLQNFATIENTPVCFSHNSLSLSLSDIDENN 1980  
 1920 GQPKILOKQSTFPQSSKDIPDRGAATDEKLQNFATIENTPVCFSHNSLSLSLSDIDENN 1979  
 1981 NKENEPINETEPPOSQOBEPSKQAGVAPKSFHVEDTPVCFSRNSSLSLSLSDIDEDDLQ 2040  
 1980 NKENEPINETEPPOSQOBEPSKQAGVAPKSFHVEDTPVCFSRNSSLSLSLSDIDEDDLQ 2039  
 2041 ECISAMPKPKKRLKGDNEKHSFRNMGGLGDEDLTLDKDIQRPDSEHGLSPDSNF 2100  
 2040 ECISAMPKPKKRLKGDNEKHSFRNMGGLGDEDLTLDKDIQRPDSEHGLSPDSNF 2099  
 2101 WKATOEGANSIVSSILHQAALACSRQASDSDSILSKGILSGSPFHLTPDOEEKPFT 2160  
 2100 WKATOEGANSIVSSILHQAALACSRQASDSDSILSKGILSGSPFHLTPDOEEKPFT 2159  
 2161 SNKGPRILLKPEKSTLETKTI ESES KGIKGGKVKYKSLITGKVESNSEI SGOMKQPIQAN 2220  
 2160 SNKGPRILLKPEKSTLETKTI ESES KGIKGGKVKYKSLITGKVESNSEI SGOMKQPIQAN 2219  
 2221 MPSISRGTMTHIFGVNRSSSTSPVSKGPPKLPATKSPSECOATTATTPRGAKPVKS 2280  
 2220 MPSISRGTMTHIFGVNRSSSTSPVSKGPPKLPATKSPSECOATTATTPRGAKPVKS 2279  
 2281 ELSPVARTQSIQIGSSKAPSRSGSRDSTPSPAPQPSRPTQSPGRNSISPRNGISPPN 2340



Db 2280 ELSPVARTSQIGSSGKAPSRGSRDSTPSRPAQPLSRPIQSPGRNSISPRNGISPPN 2339  
Qy 2341 KLSQIPRTSSPSTASTKSGGKMGYTSFGQMSQONLTQTGLSKNASSIPRSESAGK 2400  
Db 2340 KLSQIPRTSSPSTASTKSGGKMGYTSFGQMSQONLTQTGLSKNASSIPRSESAGK 2399  
Qy 2401 LNMNNGANKKVELSRMSTKSSGSESDRSERPVLRQSTFFIKEAPSPILRRKLBESA 2460  
Db 2400 LNMNNGANKKVELSRMSTKSSGSESDRSERPVLRQSTFFIKEAPSPILRRKLBESA 2459  
Qy 2461 SFESLSPSRPASPRSAQOTPVLSPLDMSLTHSSVQAGGWRKLPNLSPTIEYNDG 2520  
Db 2460 SFESLSPSRPASPRSAQOTPVLSPLDMSLTHSSVQAGGWRKLPNLSPTIEYNDG 2519  
Qy 2521 RPAKHHDARSSESPSLPINRSGTWKREHSHSSLPVSTWRTTSGSSSIISSASES 2580  
Db 2520 RPAKHHDARSSESPSLPINRSGTWKREHSHSSLPVSTWRTTSGSSSIISSASES 2579  
Qy 2581 SEKASEDEKHVNSISGTSKQKXENQVSAKGTWRKIKENEFPTNSTSTQVSSGATNGAES 2640  
Db 2580 SEKASEDEKHVNSISGTSKQKXENQVSAKGTWRKIKENEFPTNSTSTQVSSGATNGAES 2639  
Qy 2641 KTLIYOMAPAVSKTDVWVRIEDCPINNPRSGRSPGTNTPPVIDSVSEKAPNIKOSKON 2700  
Db 2640 KTLIYOMAPAVSKTDVWVRIEDCPINNPRSGRSPGTNTPPVIDSVSEKAPNIKOSKON 2699  
Qy 2701 QAKQNVGSGVPMRTVGLNRLNSFIQVADPQKGTETKPGONNPPVPSVSENESSIVERT 2760  
Db 2700 QAKQNVGSGVPMRTVGLNRLNSFIQVADPQKGTETKPGONNPPVPSVSENESSIVERT 2759  
Qy 2761 PFSSSSSKHSGSPSGTVAARVTFNNPNSPRKSSADSTSAAPSQIPTPVNNNTKRDST 2820  
Db 2760 PFSSSSSKHSGSPSGTVAARVTFNNPNSPRKSSADSTSAAPSQIPTPVNNNTKRDST 2819  
Qy 2821 DSTESSGTSQPKRHSGSVLTVTSV 2843  
Db 2820 DSTESSGTSQPKRHSGSVLTVTSV 2842

RESULT 9

US-08-289-548A-7

Sequence 7, Application US/08289546A

Patent No 5648212

GENERAL INFORMATION:

APPLICANT: ALBERTSEN, HANS

APPLICANT: ANAND, RAKESH

APPLICANT: CARLSON, MARY

APPLICANT: GRODEN, JOANNA

APPLICANT: HEDGE, PHILIP J.

APPLICANT: JOSLYN, GEORGE

APPLICANT: KINZLER, KENNETH

APPLICANT: MARKHAM, ALEXANDER F.

APPLICANT: NAKAMURA, YUSUKE

APPLICANT: THLIVERIS, ANDREW

TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC

NUMBER OF SEQUENCES: 102

CORRESPONDENCE ADDRESS:

ADDRESSEE: Banner & Allegretti, LTD

STREET: 1001 G Street, NW

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20001-4598

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/289,548A

FILING DATE: 12-AUG-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Kagan, Sarah A.

REGISTRATION NUMBER: 32,141

REFERENCE/DOCKET NUMBER: 1107.46943

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-508-9100

TELEFAX: 202-508-9299

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 2842 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

ORIGINAL SOURCE:

ORGANISM: Homo sapiens

IMMEDIATE SOURCE:

CLONE: APC

US-08-289-548A-7

Query Match 99.9%; Score 14548.5; DB 1; Length 2842;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2842; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 MAAASYDQLLKQVEALKMENSNIROBLENSNHLTKLETEASNKEVLKQLOGSIEDEAM 60  
Db 1 MAAASYDQLLKQVEALKMENSNIROBLENSNHLTKLETEASNKEVLKQLOGSIEDEAM 60  
Qy 61 ASSQIDLLRLKELNLDSSNFFGVKLRKWSLRYSVGRSGSVSSRSGECSVPVPMGSFPR 120  
Db 61 ASSQIDLLRLKELNLDSSNFFGVKLRKWSLRYSVGRSGSVSSRSGECSVPVPMGSFPR 120  
Qy 121 RGFVNGSRESTGYLEELKERSLLADLDKEEKEKWYAAQLQNLTKRIDSLPTNFSL 180  
Db 121 RGFVNGSRESTGYLEELKERSLLADLDKEEKEKWYAAQLQNLTKRIDSLPTNFSL 179  
Qy 181 QDTMTRRQLEYEARQIRVAMEEQLGTCDMEKAAORRIARICQIEKDILRIQLLOSOAT 240  
Db 180 QDTMTRRQLEYEARQIRVAMEEQLGTCDMEKAAORRIARICQIEKDILRIQLLOSOAT 239  
Qy 241 EAERSSQKHETGSHDAERQNEGQVGEINWATSGNGQSTTMDHETASVLSSSSTHSA 300  
Db 240 EAERSSQKHETGSHDAERQNEGQVGEINWATSGNGQSTTMDHETASVLSSSSTHSA 299  
Qy 301 PRRLTSHLGTKEVMYVSYLLSMLGTHDKDDMSRILLAMSSQSDSCI SMRQGCCLPLLIQLL 360  
Db 300 PRRLTSHLGTKEVMYVSYLLSMLGTHDKDDMSRILLAMSSQSDSCI SMRQGCCLPLLIQLL 359  
Qy 361 HGNDKDSVLLGNSRGSKAARASAAALHNIHSQPDQKGRREIRVILHLEQIRAYCETC 420  
Db 360 HGNDKDSVLLGNSRGSKAARASAAALHNIHSQPDQKGRREIRVILHLEQIRAYCETC 419  
Qy 421 WEQEAHEPGMDQKMPMPAPVEHQICPAVCVLMKLSFDEEHRHAMNELGGLQAIALLQ 480  
Db 420 WEQEAHEPGMDQKMPMPAPVEHQICPAVCVLMKLSFDEEHRHAMNELGGLQAIALLQ 479  
Qy 481 VDCMYGLTNDHYSITLRRYAGMALTNLTGPDVANKATLCMKMGCMRALVAQKSEEDL 540  
Db 480 VDCMYGLTNDHYSITLRRYAGMALTNLTGPDVANKATLCMKMGCMRALVAQKSEEDL 539  
Qy 541 QQVIASVLRNLSWRADVNSKKTIREVGSVKALMECALEVKKESTLKSLSALNLSAHC 600  
Db 540 QQVIASVLRNLSWRADVNSKKTIREVGSVKALMECALEVKKESTLKSLSALNLSAHC 599  
Qy 601 ENKADICAVDGAFLVGTITYSQNTNTLAIIESGGGILRNVSLLATNEDHQIILRENN 660  
Db 600 ENKADICAVDGAFLVGTITYSQNTNTLAIIESGGGILRNVSLLATNEDHQIILRENN 659  
Qy 661 CLOTLLOHLKSHSLTIIVSNACGLTNLSARNPKDQBALWDMGVSMKLNLIHSHKMIAM 720  
Db 660 CLOTLLOHLKSHSLTIIVSNACGLTNLSARNPKDQBALWDMGVSMKLNLIHSHKMIAM 719  
Qy 721 GSAAALENLNMANRPKDYKDNINSPGSSLPSLHVRKQKALEAEALDAHLSETFDNIDNLS 780



Db 720 GSAALNLANRPAKYKDANIMSPGSSPSLHVRRQKALEALDAQHLSETFDNDLNS 779  
Qy 781 PKASHRSKQHKQSLYGDYVFDNTRHDDNRSDFNTGNMTVLSPYLNTTVLPSSSSSRGS 840  
Db 780 PKASHRSKQHKQSLYGDYVFDNTRHDDNRSDFNTGNMTVLSPYLNTTVLPSSSSSRGS 839  
Qy 841 LDSRSKSLRLERERIGIGLGNTHPATENPGTSSKRGLOLSTTAAQIAKMEVSAIHTS 900  
Db 840 LDSRSKSLRLERERIGIGLGNTHPATENPGTSSKRGLOLSTTAAQIAKMEVSAIHTS 899  
Qy 901 QEDRSSGTTTELHCVTDERNALRRSSAAHSTNYNFTKSENNRCTSMPIAKLEYKRSS 960  
Db 900 QEDRSSGTTTELHCVTDERNALRRSSAAHSTNYNFTKSENNRCTSMPIAKLEYKRSS 959  
Qy 961 NDSLNSVSSDGYGKRGOMKPSIESYSEDDESKFCYGOYPADLAHKIHSANHMDNDGE 1020  
Db 960 NDSLNSVSSDGYGKRGOMKPSIESYSEDDESKFCYGOYPADLAHKIHSANHMDNDGE 1019  
Qy 1021 LDTPINYLKYSDBOLNSGRQSPSONERWARPKHITTEDEIKQSEORQSRNOSTTYPVYTE 1080  
Db 1020 LDTPINYLKYSDBOLNSGRQSPSONERWARPKHITTEDEIKQSEORQSRNOSTTYPVYTE 1079  
Qy 1081 STDDKHLKFPHFQOCVCVPSYRSGANGSETNRVGSNHNQVNSQSCQEDDYEDDKP 1140  
Db 1080 STDDKHLKFPHFQOCVCVPSYRSGANGSETNRVGSNHNQVNSQSCQEDDYEDDKP 1139  
Qy 1141 TNSYRYSEREBQHEERPTNYSIKYNEEKRHVDQPIDYSLKYATDIPSSQKQSPFSKS 1200  
Db 1140 TNSYRYSEREBQHEERPTNYSIKYNEEKRHVDQPIDYSLKYATDIPSSQKQSPFSKS 1199  
Qy 1201 SSGQSKTEHWSSENSTPSSNAKQNLHPSSAQSRGQPKAATCKVSSINQETIQ 1260  
Db 1200 SSGQSKTEHWSSENSTPSSNAKQNLHPSSAQSRGQPKAATCKVSSINQETIQ 1259  
Qy 1261 TYCVEDTPICFRCSLSLSSAEDDEIGCQNTQOEAASANTLOIAIEKIKIGTRSAEDPV 1320  
Db 1260 TYCVEDTPICFRCSLSLSSAEDDEIGCQNTQOEAASANTLOIAIEKIKIGTRSAEDPV 1319  
Qy 1321 SEVPAVQSHERTKSLRQGLSSLSAESAARHKAVFSSGAKSPSKGACTPKSPPEHVQET 1380  
Db 1320 SEVPAVQSHERTKSLRQGLSSLSAESAARHKAVFSSGAKSPSKGACTPKSPPEHVQET 1379  
Qy 1381 PLMFSRCTSVSSLDSPESRSIASSVQSEPCSGMVSGIISPSDLPDPGQTMPPSRKTPP 1440  
Db 1380 PLMFSRCTSVSSLDSPESRSIASSVQSEPCSGMVSGIISPSDLPDPGQTMPPSRKTPP 1439  
Qy 1441 PPQTAQTKREVPKNKAFTAKRESGPKQAANVAQVQVLPDADTLHFAESTPDGF 1500  
Db 1440 PPQTAQTKREVPKNKAFTAKRESGPKQAANVAQVQVLPDADTLHFAESTPDGF 1499  
Qy 1501 SCSSLSALSILDEPFIQKVELRIMPVQENDNGNETESEQPKESNENQKAEKTIIDSE 1560  
Db 1500 SCSSLSALSILDEPFIQKVELRIMPVQENDNGNETESEQPKESNENQKAEKTIIDSE 1559  
Qy 1561 KDLLDDSDDDDEILEBECIISAMPTKSSRKAQPAQATASKLPPPVARKPSQLPVYKLLPS 1620  
Db 1560 KDLLDDSDDDDEILEBECIISAMPTKSSRKAQPAQATASKLPPPVARKPSQLPVYKLLPS 1619  
Qy 1621 QNRLOPKHVSFTPGDDMPRVYCVEGTPIFNSTATSLSDLTIESPPNELAAGEVGRVGAQ 1680  
Db 1620 QNRLOPKHVSFTPGDDMPRVYCVEGTPIFNSTATSLSDLTIESPPNELAAGEVGRVGAQ 1679  
Qy 1681 SGBFEKEDTPTTEGRSTDEAQGKTGSVTIPELDDNKAEEGDILABEINSAHPKGSHP 1740  
Db 1680 SGBFEKEDTPTTEGRSTDEAQGKTGSVTIPELDDNKAEEGDILABEINSAHPKGSHP 1739  
Qy 1741 FRVKKINDVQQAASASSAPKNQOLGKKKPTSPVKPIPONTYETRVRKNADSKNNLN 1800  
Db 1740 FRVKKINDVQQAASASSAPKNQOLGKKKPTSPVKPIPONTYETRVRKNADSKNNLN 1799  
Qy 1801 AERVFSNKKSKQNLKNNKDFNDKLPNNEDRVGRSFAPDSPHHTYPIEGTPYCFSRND 1860

Db 1800 AERVFSNKKSKQNLKNNKDFNDKLPNNEDRVGRSFAPDSPHHTYPIEGTPYCFSRND 1859  
Qy 1861 SILSLDFDDDDVDLSREKAELRKAKENKESAKVTSHTELTSNQOSANKTOIAIAKQPINR 1920  
Db 1860 SILSLDFDDDDVDLSREKAELRKAKENKESAKVTSHTELTSNQOSANKTOIAIAKQPINR 1919  
Qy 1921 COPRPILOKQSTFPQSSKDIIPDRGAATDEKLOFAIENTPVCFSHNSLSLSLSDIDQENN 1980  
Db 1920 COPRPILOKQSTFPQSSKDIIPDRGAATDEKLOFAIENTPVCFSHNSLSLSLSDIDQENN 1979  
Qy 1981 NKENEPKETETPPDOSQBPSPQASGVAPKSFHVEDTPVCFSRNSSLSLSLSDIDEDLLQ 2040  
Db 1980 NKENEPKETETPPDOSQBPSPQASGVAPKSFHVEDTPVCFSRNSSLSLSLSDIDEDLLQ 2039  
Qy 2041 ECISAMPKKKPRLKGDNEKHSFRMGGILGDEDLTLDKDIQRPDSEHGLSPDSENF 2100  
Db 2040 ECISAMPKKKPRLKGDNEKHSFRMGGILGDEDLTLDKDIQRPDSEHGLSPDSENF 2099  
Qy 2101 WKAIQEGANGSIVSSLHOAAAAACLSROASSDSDSILSKGISLGSPPHLPDQEEKPFT 2160  
Db 2100 WKAIQEGANGSIVSSLHOAAAAACLSROASSDSDSILSKGISLGSPPHLPDQEEKPFT 2159  
Qy 2161 SNKGPRILKPGESTLETETKIESESKGIGKGGKYYKSLITGKVRNSSEISGQMKQPLQAN 2220  
Db 2160 SNKGPRILKPGESTLETETKIESESKGIGKGGKYYKSLITGKVRNSSEISGQMKQPLQAN 2219  
Qy 2221 MPSISRGRTMIHPGVNRNSSSTSPVSKGPPPLTPASKSPSEGTATTPRGAKPSPKS 2280  
Db 2220 MPSISRGRTMIHPGVNRNSSSTSPVSKGPPPLTPASKSPSEGTATTPRGAKPSPKS 2279  
Qy 2281 ELSPVARTSIOGSSKAPSRSGSDSTPSRPAQPLSRPIQSPGRNSISPGRNGISPPN 2340  
Db 2280 ELSPVARTSIOGSSKAPSRSGSDSTPSRPAQPLSRPIQSPGRNSISPGRNGISPPN 2339  
Qy 2341 KLSQLPRTSSTASTKSSGSGKMSYTSRQMSQOQLTKQTGLSKVASSIPRSESASK 2400  
Db 2340 KLSQLPRTSSTASTKSSGSGKMSYTSRQMSQOQLTKQTGLSKVASSIPRSESASK 2399  
Qy 2401 LNQMNGANGANKVELSRMSTKSSGESRDERPVLVROSTFIKEAPSTLRRKLBSA 2460  
Db 2400 LNQMNGANGANKVELSRMSTKSSGESRDERPVLVROSTFIKEAPSTLRRKLBSA 2459  
Qy 2461 SFESLSPSRSPASTRQAQTPVLPSLPMDSLTHSSVQAGWRKLPNLSPTIEYNDG 2520  
Db 2460 SFESLSPSRSPASTRQAQTPVLPSLPMDSLTHSSVQAGWRKLPNLSPTIEYNDG 2519  
Qy 2521 RPAKHDIARSHSSPSRLPINRSGTWKREHSKSSSLPRVSTWRRRTGSSSSILSASSES 2580  
Db 2520 RPAKHDIARSHSSPSRLPINRSGTWKREHSKSSSLPRVSTWRRRTGSSSSILSASSES 2579  
Qy 2581 SEKAKSEDEKHVNSISGTQSKENQVSAKGTWRKIKENEFSPNTNSTQTVSSGATNGAES 2640  
Db 2580 SEKAKSEDEKHVNSISGTQSKENQVSAKGTWRKIKENEFSPNTNSTQTVSSGATNGAES 2639  
Qy 2641 KTLIYQMAPASKTEDYWRLEDPCINNRSGRSPGTNTPTPVDVSEKANPNIKSKDN 2700  
Db 2640 KTLIYQMAPASKTEDYWRLEDPCINNRSGRSPGTNTPTPVDVSEKANPNIKSKDN 2699  
Qy 2701 QAKNVNGSVPMRTVGLNRLNSFIQVADPDQKGTETIKPGQNNPVPVSETNESSIVERT 2760  
Db 2700 QAKNVNGSVPMRTVGLNRLNSFIQVADPDQKGTETIKPGQNNPVPVSETNESSIVERT 2759  
Qy 2761 PFSSSSSKHSSPSGTVAAARVTPPNYNPSPRKSSADTSARPQIPTPVNNNTKKRDSKT 2820  
Db 2760 PFSSSSSKHSSPSGTVAAARVTPPNYNPSPRKSSADTSARPQIPTPVNNNTKKRDSKT 2819  
Qy 2821 DSTESSGTQSPKHSGLVLTISV 2843  
Db 2820 DSTESSGTQSPKHSGLVLTISV 2842

Sequence 7, Application US/08452654  
Patent No. 5691454  
GENERAL INFORMATION:  
APPLICANT: ALBERTSEN, HANS  
APPLICANT: ANAND, RAKESH  
APPLICANT: CARLSON, MARY  
APPLICANT: GRODEN, JOANNA  
APPLICANT: HEDGE, PHILIP J.  
APPLICANT: JOSLYN, GEOFF  
APPLICANT: KINZLER, KENNETH  
APPLICANT: NAKHAM, ALEXANDER F.  
APPLICANT: NAKAMURA, YUSUKE  
APPLICANT: THLIVERIS, ANDREW  
TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC  
TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS  
NUMBER OF SEQUENCES: 94  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Banner, Birch, McKie & Beckett  
STREET: 1001 G Street, NW  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20001-4598  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/452,654  
FILING DATE: 25-MAY-1995  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/741,940  
FILING DATE: 08-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Kagan, Sarah A.  
REGISTRATION NUMBER: 32,141  
REFERENCE/DOCKET NUMBER: 1107.035574  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-508-9100  
TELEFAX: 202-508-9299  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2842 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
IMMEDIATE SOURCE:  
CLONE: APC  
US-08-452-654-7

Query Match 99.9%; Score 14548.5; DB 1; Length 2842;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2842; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MAASVDQLLKVEALKMNSNLRQLEDNSNHLTKLETSANMKEVLKQLQGSIEDAM 60  
DB 1 MAASVDQLLKVEALKMNSNLRQLEDNSNHLTKLETSANMKEVLKQLQGSIEDAM 60  
QY 61 ASSGQIDLLERLKEALNLDSSNFPFVKLRSMKSLRSGREGSVSRSGECSPVPMGSPFR 120  
DB 61 ASSGQIDLLERLKEALNLDSSNFPFVKLRSMKSLRSGREGSVSRSGECSPVPMGSPFR 120  
QY 121 RGFVNGSRESTGYLBELEKERSILLADLDKEKEKDWYQAQNLTKEDISLPIFENESL 180  
DB 121 RGFVNGSRESTGYLBELEKERSILLADLDKEKEKDWYQAQNLTKEDISLPIFENESL 179  
QY 181 QDTMTTRQLEYEARQIRVAMEEQLGTCQDMKEKAQRRIARIQOIEKDIILIRQLLSQAT 240

180 QDTMTTRQLEYEARQIRVAMEEQLGTCQDMKEKAQRRIARIQOIEKDIILIRQLLSQAT 239  
241 EAERSSQKHETGSHDAERQNEGGQYGEINMATSNGQGGSTTRMDHETASVLSSSTHSA 300  
240 EAERSSQKHETGSHDAERQNEGGQYGEINMATSNGQGGSTTRMDHETASVLSSSTHSA 299  
301 PRRLTSHLGTKEVMYVSLLSMLGTHDKDDMSRLLAMSSQDSCISMRSGCPLLIQLL 360  
300 PRRLTSHLGTKEVMYVSLLSMLGTHDKDDMSRLLAMSSQDSCISMRSGCPLLIQLL 359  
361 HGNDKDSVLLGNSRSGKEARASAAALHNIHSQPDQDKRGRREIRVJLHLEQIRAYCETC 420  
360 HGNDKDSVLLGNSRSGKEARASAAALHNIHSQPDQDKRGRREIRVJLHLEQIRAYCETC 419  
421 NEWQEAHEPGMDQDKNPMFAPVEHQICPAVCULMKLSFDEEHRHANVELGGQAIALLQ 480  
420 NEWQEAHEPGMDQDKNPMFAPVEHQICPAVCULMKLSFDEEHRHANVELGGQAIALLQ 479  
481 VDCEMYGLTNDHYSITLRRYAGMALTNLTFDGVANKATLCSMKGCMRALVAQLKSSSEDL 540  
480 VDCEMYGLTNDHYSITLRRYAGMALTNLTFDGVANKATLCSMKGCMRALVAQLKSSSEDL 539  
541 QQVIASVLRNLSWRADVNSKTLREVGSKVAMECALEVKKESTLKSVALNLSAHCT 600  
540 QQVIASVLRNLSWRADVNSKTLREVGSKVAMECALEVKKESTLKSVALNLSAHCT 599  
601 ENKADICAVDGAFLVGLTYESQNTLAIIESGGGILRNVSLLIATNEDHRCIILRENN 660  
600 ENKADICAVDGAFLVGLTYESQNTLAIIESGGGILRNVSLLIATNEDHRCIILRENN 659  
661 CLQTLQLHLKSHSLTIVSNACGLTNLSARNPKDQBALMDMGAVSMLKNLIHSHKHMIAM 720  
660 CLQTLQLHLKSHSLTIVSNACGLTNLSARNPKDQBALMDMGAVSMLKNLIHSHKHMIAM 719  
721 GSAAALRNLMANRPKAKYKDANINSPGSLPSLHVKKOKALEAELDAHLSEFNDNLNLS 780  
720 GSAAALRNLMANRPKAKYKDANINSPGSLPSLHVKKOKALEAELDAHLSEFNDNLNLS 779  
781 PKASHRSKQKHQSLYGDYVFDNRRHDDNRSDNFNTGNMTVLSPLYNTTLVLPSSSSSRGS 840  
780 PKASHRSKQKHQSLYGDYVFDNRRHDDNRSDNFNTGNMTVLSPLYNTTLVLPSSSSSRGS 839  
841 LDSRSSEKORSLERRGIGLGNTHPATENPGTSKKGLOISTTAAQIAKMEVSVSAIHTS 900  
840 LDSRSSEKORSLERRGIGLGNTHPATENPGTSKKGLOISTTAAQIAKMEVSVSAIHTS 899  
901 QEDRSSGTTTELHCVTDERNALRSAAATHSNFTKSENSNRITCSMPYAKLEYKRSS 960  
900 QEDRSSGTTTELHCVTDERNALRSAAATHSNFTKSENSNRITCSMPYAKLEYKRSS 959  
961 NDSLNSVSSSDGYGKRGQMKPSIESYSEDESFCYGYQYPADLAHKIHSANHMDNDGE 1020  
960 NDSLNSVSSSDGYGKRGQMKPSIESYSEDESFCYGYQYPADLAHKIHSANHMDNDGE 1019  
1021 LDTPIYSLKYDEQLNSGRQSPSQNERWARPKHIIIEDEIKOSEQORSNQSTTYVYTE 1080  
1020 LDTPIYSLKYDEQLNSGRQSPSQNERWARPKHIIIEDEIKOSEQORSNQSTTYVYTE 1079  
1081 STDDKHLKFQPHFGQCEVSPYRSRGANGSETNRVNGSHNGINQVNSQSCQEDDYDDKP 1140  
1080 STDDKHLKFQPHFGQCEVSPYRSRGANGSETNRVNGSHNGINQVNSQSCQEDDYDDKP 1139  
1141 TNYSERYSBEOHEBERPTNYSIKYNEEKHVDQPIDYSLKYATDIPSSQKGSFSPKS 1200  
1140 TNYSERYSBEOHEBERPTNYSIKYNEEKHVDQPIDYSLKYATDIPSSQKGSFSPKS 1199  
1201 SSGCSSKTEHMSSSSENTSTPSSNAKRONLHPSSAQSRSGQPKAATCKVSSINQETIQ 1260  
1200 SSGCSSKTEHMSSSSENTSTPSSNAKRONLHPSSAQSRSGQPKAATCKVSSINQETIQ 1259  
1261 TYCVEDTPICFSRCSSISLSSAEDIGCNQTTQEADSAANTLQIAIIEKIGITRSEADPV 1320  
1260 TYCVEDTPICFSRCSSISLSSAEDIGCNQTTQEADSAANTLQIAIIEKIGITRSEADPV 1319

1321 SEVPVAVQHPRTKSSRIQGSLSSEGAHKAFFSSGAKSPSKGAQTPKSPPEHYVQET 1380  
 1320 SEVPVAVQHPRTKSSRIQGSLSSEGAHKAFFSSGAKSPSKGAQTPKSPPEHYVQET 1379  
 1381 PLMFSTRCTSVSSLDSPESRISASVQSPGCMVSGIISPSDLDPSPGOTMPPSRSKTPP 1440  
 1380 PLMFSTRCTSVSSLDSPESRISASVQSPGCMVSGIISPSDLDPSPGOTMPPSRSKTPP 1439  
 1441 PPQTAQTKREVKNKAPTAKEKESGPKQAANVAQVQVLPDADTLHLHFAESTPDGF 1500  
 1440 PPQTAQTKREVKNKAPTAKEKESGPKQAANVAQVQVLPDADTLHLHFAESTPDGF 1499  
 1501 SCSSLSALSALDPFFIQKQVLELIMPPVQENDNGNETSEPOKESNENQEKAEKTI DSE 1560  
 1500 SCSSLSALSALDPFFIQKQVLELIMPPVQENDNGNETSEPOKESNENQEKAEKTI DSE 1559  
 1561 KDLLDDSDDDDDIIEILBECIISAMFTKSSRKAKKPAQTAASKLPPPVARKPSQLPVYKLLPS 1620  
 1560 KDLLDDSDDDDDIIEILBECIISAMFTKSSRKAKKPAQTAASKLPPPVARKPSQLPVYKLLPS 1619  
 1621 QNLQPKQVSVFTPGDDMPRVYCEGTPINFSTATSLDLTIBSPNNEAAGEVGRGGAQ 1680  
 1620 QNLQPKQVSVFTPGDDMPRVYCEGTPINFSTATSLDLTIBSPNNEAAGEVGRGGAQ 1679  
 1681 SGFEKRDITPTBGRSTDEAQQGKTSVITPELDDNKABEGDILACINSAMPKGSHPK 1740  
 1680 SGFEKRDITPTBGRSTDEAQQGKTSVITPELDDNKABEGDILACINSAMPKGSHPK 1739  
 1741 FRVKIMQVQOASASSAPKNQOLGKKKTPSPVKPIQNTTEYTRVRKNADSKNIN 1800  
 1740 FRVKIMQVQOASASSAPKNQOLGKKKTPSPVKPIQNTTEYTRVRKNADSKNIN 1799  
 1801 AERFVSNKDKSKONLKNNSKDFNDKLPNNEDVRGSEFADSPHHVTPTEGGPYCFERN 1860  
 1800 AERFVSNKDKSKONLKNNSKDFNDKLPNNEDVRGSEFADSPHHVTPTEGGPYCFERN 1859  
 1861 SLSSLDFFDDDDVLSREKAELEKAKENKSEAKVTSHTELTNSQOSANKTQAIKQPINR 1920  
 1860 SLSSLDFFDDDDVLSREKAELEKAKENKSEAKVTSHTELTNSQOSANKTQAIKQPINR 1919  
 1921 GQPKPILQKOSTPQSSKOTPDRCATDEKLONFATENTPVCFSHNSLSLSIDDOENN 1980  
 1920 GQPKPILQKOSTPQSSKOTPDRCATDEKLONFATENTPVCFSHNSLSLSIDDOENN 1979  
 1981 NKNEPIKETEPDPSQOEPSPKQASGYAPKSFHVEDTPVCFERNSSLSLSIDSDDLQ 2040  
 1980 NKNEPIKETEPDPSQOEPSPKQASGYAPKSFHVEDTPVCFERNSSLSLSIDSDDLQ 2039  
 2041 ECISAMPKPKKPSRLKGDNEKHSRNMGGIIGEDLTLDKQIRPDPSHGLSPGSENF 2100  
 2040 ECISAMPKPKKPSRLKGDNEKHSRNMGGIIGEDLTLDKQIRPDPSHGLSPGSENF 2099  
 2101 WKAIQEGANSIVSLHQAASAAALCLQASSDSDSILSLKSGISLGSPFFHLTPDQEKPT 2160  
 2100 WKAIQEGANSIVSLHQAASAAALCLQASSDSDSILSLKSGISLGSPFFHLTPDQEKPT 2159  
 2161 SNKGPRILKPEKSTLETTKKIESEKGIKGGKVKYKSLITGKVRNSSEISGQWKQPLQAN 2220  
 2160 SNKGPRILKPEKSTLETTKKIESEKGIKGGKVKYKSLITGKVRNSSEISGQWKQPLQAN 2219  
 2221 MPISRGRTMIHLPVGNSSSTSPVSKGPPKLTPTASKSPBEGOTATTSPRGAKPSVKS 2280  
 2220 MPISRGRTMIHLPVGNSSSTSPVSKGPPKLTPTASKSPBEGOTATTSPRGAKPSVKS 2279  
 2281 ELSFVARQTSIQGSSKAPSRGSRDSTPSRPAQQLSRPIQSPGRNISPGRNGISPPN 2340  
 2280 ELSFVARQTSIQGSSKAPSRGSRDSTPSRPAQQLSRPIQSPGRNISPGRNGISPPN 2339  
 2341 KLSQLPRTSPSTASTKSSGSGMSYTSRGRQNSQNLTKQGLSKNASSIPRSESASKG 2400  
 2340 KLSQLPRTSPSTASTKSSGSGMSYTSRGRQNSQNLTKQGLSKNASSIPRSESASKG 2399

2401 LNQMNGNANKVELSRMSTKSSGSESDRSPVILVROSTFIKEAPSPTLRKLKLESA 2460  
 2400 LNQMNGNANKVELSRMSTKSSGSESDRSPVILVROSTFIKEAPSPTLRKLKLESA 2459  
 2461 SFESLSPSSRRPASPTRSQAOQTPVLSPLPDMSLSTHSSVQAGWRKLPPLSPPTIEYNDG 2520  
 2460 SFESLSPSSRRPASPTRSQAOQTPVLSPLPDMSLSTHSSVQAGWRKLPPLSPPTIEYNDG 2519  
 2521 RPAKHDIARSHSPSRPLPINRSGTWKREHKSUSLSIPRVSTWRTGSSSSILSASSES 2580  
 2520 RPAKHDIARSHSPSRPLPINRSGTWKREHKSUSLSIPRVSTWRTGSSSSILSASSES 2579  
 2581 SEKAKSEDEKHVNSISGTKQSKENQVSAKGTWRIKENEPSPTNSTQTVSSGATNGAES 2640  
 2580 SEKAKSEDEKHVNSISGTKQSKENQVSAKGTWRIKENEPSPTNSTQTVSSGATNGAES 2639  
 2641 KTLIYQMAPAVSKTEDVWVRIEDCPINNPRSGRPTGNTPPVIDSVSEKANPNIKDSKN 2700  
 2640 KTLIYQMAPAVSKTEDVWVRIEDCPINNPRSGRPTGNTPPVIDSVSEKANPNIKDSKN 2699  
 2701 QAKQNVGNGSVPMETVGLNELNSFIQVADPDQKGTIKQCONNPVSVSETNESSIVERT 2760  
 2700 QAKQNVGNGSVPMETVGLNELNSFIQVADPDQKGTIKQCONNPVSVSETNESSIVERT 2759  
 2761 PFSSSSSKHSPSGTVAARVTPFNYPNPSPKSSADSTSPARPSQIPTPVNNNTKKRDSKT 2820  
 2760 PFSSSSSKHSPSGTVAARVTPFNYPNPSPKSSADSTSPARPSQIPTPVNNNTKKRDSKT 2819  
 2821 DSTESSGTQSPKHSGLVLTSV 2843  
 2820 DSTESSGTQSPKHSGLVLTSV 2842

## RESULT 11

US-08-443-731-7

; Sequence 7, Application US/08449731

; Patent No. 6413727

; GENERAL INFORMATION:

; APPLICANT: ALBERTSEN, HANS

; ANAND, RAKESH

; CARLSON, MARY

; GRODEN, JOANNA

; HEDGE, PHILIP J.

; JOSLYN, GEOFF

; KINZLER, KENNETH

; MARKHAM, ALEXANDER F.

; NAKAMURA, YUSUKE

; THLIVERIS, ANDREW

; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC

; GENE IN COLORECTAL CANCER IN HUMANS

; NUMBER OF SEQUENCES: 102

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Banner &amp; Allegretti, LTD

; STREET: 1001 G Street, NW

; CITY: Washington

; STATE: D.C.

; COUNTRY: USA

; ZIP: 20001-4598

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/449,731

; FILING DATE: 25-May-1995

; CLASSIFICATION: &lt;Unknown&gt;

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: 08/289,548

; FILING DATE: 12-AUG-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Kagan, Sarah A.

; REGISTRATION NUMBER: 32,141

REFERENCE/DOCKET NUMBER: 1107.46943  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-508-9100  
TELEFAX: 202-508-9299  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2842 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
IMMEDIATE SOURCE:  
CLONE: APC  
SEQUENCE DESCRIPTION: SEQ ID NO: 7:  
US-08-449-731-7

Query Match 99.9%; Score 14548.5; DB 4; Length 2842;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2842; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY	1	MAAASVDQLLKQVEALKVENSNIROLEDNSNHLTKLETEASNKEVLKQCGSTIEDAM	60
DB	1	MAAASVDQLLKQVEALKVENSNIROLEDNSNHLTKLETEASNKEVLKQCGSTIEDAM	60
QY	61	ASSGGIDLLERLKEINLNDSSNPPGVKLRSKMSLSRSGREGSVSRSGCECPVPMGSPFR	120
DB	61	ASSGGIDLLERLKEINLNDSSNPPGVKLRSKMSLSRSGREGSVSRSGCECPVPMGSPFR	120
QY	121	RGFVNGSRSTGYLSELEKERSLLADLDKBEKXDWYIAQONLTKRIDSPLTENPSL	180
DB	121	RGFVNGSRSTGYLSELEKERSLLADLDKBEKXDWYIAQONLTKRIDSPLTENPSL	179
QY	181	QDTMTROLEYEAROIRVAMEBOLGTCODMEKRAORRIARTQOEKDIRLRLQLOSOAT	240
DB	180	QDTMTROLEYEAROIRVAMEBOLGTCODMEKRAORRIARTQOEKDIRLRLQLOSOAT	239
QY	241	EAERSSONKHETGSHDAERQEGVGGEINNATSGNGQGTTRMDHETASVLSSTHSA	300
DB	240	EAERSSONKHETGSHDAERQEGVGGEINNATSGNGQGTTRMDHETASVLSSTHSA	299
QY	301	PRRLTSHLGTKVEMYVLSLLSMGLTHDKDDMSHTILAMSSODSCISMEQSGCLPLIQLL	360
DB	300	PRRLTSHLGTKVEMYVLSLLSMGLTHDKDDMSHTILAMSSODSCISMEQSGCLPLIQLL	359
QY	361	HGNDXDSVLLGNSRGSKEARAFASAAALNIIHSQPDCKRGRREIRVLHLLQIRAYCETC	420
DB	360	HGNDXDSVLLGNSRGSKEARAFASAAALNIIHSQPDCKRGRREIRVLHLLQIRAYCETC	419
QY	421	NEWQBAEPFGMDQDQNPAPVVEHOICPAVCVLMKLSFDEBHRHAMNELGGLQIAELLQ	480
DB	420	NEWQBAEPFGMDQDQNPAPVVEHOICPAVCVLMKLSFDEBHRHAMNELGGLQIAELLQ	479
QY	481	VDCEMYGLTNDHYSITLRRYAGMALTNLTFGDVANKATLCMKGCMRALVAQLKSEBDL	540
DB	480	VDCEMYGLTNDHYSITLRRYAGMALTNLTFGDVANKATLCMKGCMRALVAQLKSEBDL	539
QY	541	QOVIASVLNLSWRADVNSKTLRLRVGSVKALMECALEVYKKESTLKSVLNALWLSAHC	600
DB	540	QOVIASVLNLSWRADVNSKTLRLRVGSVKALMECALEVYKKESTLKSVLNALWLSAHC	599
QY	601	ENKADICAVDGGALAFVLGTLTVRSQNTLAIIESGGILRVNSSLIATNEHQRLRENN	660
DB	600	ENKADICAVDGGALAFVLGTLTVRSQNTLAIIESGGILRVNSSLIATNEHQRLRENN	659
QY	661	CIQTLQLHLKSHSLTVSNACGTLNLSARNPKDQEAALWDMGAVSMKLIHSHKMIAM	720
DB	660	CIQTLQLHLKSHSLTVSNACGTLNLSARNPKDQEAALWDMGAVSMKLIHSHKMIAM	719
QY	721	GSAALRNLMANRPAPYKXDNIMTSGSSLPSSHVRKQKALEABELDAQHLSLTFNIDNLS	780
DB	720	GSAALRNLMANRPAPYKXDNIMTSGSSLPSSHVRKQKALEABELDAQHLSLTFNIDNLS	779

QY	781	PKASHRSKQHKOSLKYGDYVFDNRHDDNRSDNFTGNMTVLSPYLNTTVLPSSSSSRGS	840
DB	780	PKASHRSKQHKOSLKYGDYVFDNRHDDNRSDNFTGNMTVLSPYLNTTVLPSSSSSRGS	839
QY	841	LDSSRSSEKDRSLERERIGLGNYPATENPGTSSKEGLQISTTAAQIAKVMEEVSAIHTS	900
DB	840	LDSSRSSEKDRSLERERIGLGNYPATENPGTSSKEGLQISTTAAQIAKVMEEVSAIHTS	899
QY	901	QEDRSSGTTTELHCVTDERNALRRSSAAATHSYNTYFTKSENENRTCSMPYAKLEYKRSS	960
DB	900	QEDRSSGTTTELHCVTDERNALRRSSAAATHSYNTYFTKSENENRTCSMPYAKLEYKRSS	959
QY	961	NDLSNVSSSDGYGKRGOMKPSIESYSEDDSEKFCYGYQVPADLAHKIHSANEMDNDGE	1020
DB	960	NDLSNVSSSDGYGKRGOMKPSIESYSEDDSEKFCYGYQVPADLAHKIHSANEMDNDGE	1019
QY	1021	LDPINYSKYSEQLNSGRQSPSONERWAPKHIIIEDEIKQSEQRNQSTTYPVYTE	1080
DB	1020	LDPINYSKYSEQLNSGRQSPSONERWAPKHIIIEDEIKQSEQRNQSTTYPVYTE	1079
QY	1081	STDDKHLKQFHFGQCECVSPYSRGANGSETNRVGSNHGINONVQSLOCEDDYEDDKP	1140
DB	1080	STDDKHLKQFHFGQCECVSPYSRGANGSETNRVGSNHGINONVQSLOCEDDYEDDKP	1139
QY	1141	TNYSERYSEEEQHEEERPTNYSIKYNEEKRHVDQPIDYSLKYATDIPSSQKQSFSPKS	1200
DB	1140	TNYSERYSEEEQHEEERPTNYSIKYNEEKRHVDQPIDYSLKYATDIPSSQKQSFSPKS	1199
QY	1201	SSQSSKTEHMSSENSTPSSNAKRONQLHPSSAQSRSGQPKAATCKVSSINQETIQ	1260
DB	1200	SSQSSKTEHMSSENSTPSSNAKRONQLHPSSAQSRSGQPKAATCKVSSINQETIQ	1259
QY	1261	TYCVEDTPICFSSCSLSLSAEDIGCNOTTQEADSNANTLQIAIKIKIGTRSAEDPV	1320
DB	1260	TYCVEDTPICFSSCSLSLSAEDIGCNOTTQEADSNANTLQIAIKIKIGTRSAEDPV	1319
QY	1321	SEVPASVQHPRTKSRLQSSLSSESARHKAVFSSGAKSPSKGACTPKGPEHYVQET	1380
DB	1320	SEVPASVQHPRTKSRLQSSLSSESARHKAVFSSGAKSPSKGACTPKGPEHYVQET	1379
QY	1381	PLMFSRCTSVSSLDSPESRSIASVQSEPCSGMVSGIISPSDLPDGPQTMPSPRSKTPP	1440
DB	1380	PLMFSRCTSVSSLDSPESRSIASVQSEPCSGMVSGIISPSDLPDGPQTMPSPRSKTPP	1439
QY	1441	PPQTACTKREVPKNKAPTAARESGPKQAAVNAVQVQVLPDADTLLHFATESTPDGF	1500
DB	1440	PPQTACTKREVPKNKAPTAARESGPKQAAVNAVQVQVLPDADTLLHFATESTPDGF	1499
QY	1501	SCSSLSALSIDBFFIQKDYELRIMPVQENDNGNETESEOPKESNENQEKAEKTIIDSE	1560
DB	1500	SCSSLSALSIDBFFIQKDYELRIMPVQENDNGNETESEOPKESNENQEKAEKTIIDSE	1559
QY	1561	KDLDSDSDDDIIELEECIIISAMPTKSSRKAKKPAOTASKLPPPVARKPSQLPVYKLLPS	1620
DB	1560	KDLDSDSDDDIIELEECIIISAMPTKSSRKAKKPAOTASKLPPPVARKPSQLPVYKLLPS	1619
QY	1621	QNRLOPKHVSFTPGDDMPRVYCVGEGTPIINFSTATSLSDLTIESPPNELAAGVGVRGAQ	1680
DB	1620	QNRLOPKHVSFTPGDDMPRVYCVGEGTPIINFSTATSLSDLTIESPPNELAAGVGVRGAQ	1679
QY	1681	SGFEKEDTITPTGRSTDEAQQGKTSSVTPELDDNKAEBEGDILAEICINSAMPKGSHKP	1740
DB	1680	SGFEKEDTITPTGRSTDEAQQGKTSSVTPELDDNKAEBEGDILAEICINSAMPKGSHKP	1739
QY	1741	FRVKIMDQVQOASASSAPNKNQDKKKKPTSPVKPIPQNTTEYTRVRKNADSKNNLN	1800
DB	1740	FRVKIMDQVQOASASSAPNKNQDKKKKPTSPVKPIPQNTTEYTRVRKNADSKNNLN	1799
QY	1801	AERFSDNKDKKONLKNNSKDFNDKLPNNEDSVRGSAFDSPHHYTPIEGTTCYCSRND	1860
DB	1800	AERFSDNKDKKONLKNNSKDFNDKLPNNEDSVRGSAFDSPHHYTPIEGTTCYCSRND	1859

```

QY 1861 SLGLDDEDDVLSRKAELRKAENKESAKVTSHTLTSTNOQANKTQAIKQPINR 1920
DB 1860 SLGLDDEDDVLSRKAELRKAENKESAKVTSHTLTSTNOQANKTQAIKQPINR 1919
QY 1921 GQPKILKOSTTPOSSKDI PDGAATDEKLQNFALIENTVCFSHNSSLSLSDIDQENN 1980
DB 1920 GQPKILKOSTTPOSSKDI PDGAATDEKLQNFALIENTVCFSHNSSLSLSDIDQENN 1979
QY 1981 NKNEPIKETEPDQSGEPKQASGYAPKSFVETDTPVCFSRNSSLSLSDIDQENLLQ 2040
DB 1980 NKNEPIKETEPDQSGEPKQASGYAPKSFVETDTPVCFSRNSSLSLSDIDQENLLQ 2039
QY 2041 ECISAPKPKKPSRLKGDNEKHSPRNMGILGEDTLDKDQRPDSEHGLSPDSENF 2100
DB 2040 ECISAPKPKKPSRLKGDNEKHSPRNMGILGEDTLDKDQRPDSEHGLSPDSENF 2099
QY 2101 WKATQEGANSIVSLHQAACLSRQASDSDSILSLKSGISLGSPFHLTPDQEEKPFT 2160
DB 2100 WKATQEGANSIVSLHQAACLSRQASDSDSILSLKSGISLGSPFHLTPDQEEKPFT 2159
QY 2161 SNKGPRILKPEKSTLTETKIEBSKIGKGGKVKYKSLITGKVRNSSEISGQWKQPLQAN 2220
DB 2160 SNKGPRILKPEKSTLTETKIEBSKIGKGGKVKYKSLITGKVRNSSEISGQWKQPLQAN 2219
QY 2221 MPSISRGRTWIHPGVNNSSTSPVSKGPPPLKTPASKPSSEGQTATTSRGAKPVSXK 2280
DB 2220 MPSISRGRTWIHPGVNNSSTSPVSKGPPPLKTPASKPSSEGQTATTSRGAKPVSXK 2279
QY 2281 ELSVPARTSQIGSSKAPSRGSRDSTPGRPAQPLSRPIQSPGRNSISPGRNGISPPN 2340
DB 2280 ELSVPARTSQIGSSKAPSRGSRDSTPGRPAQPLSRPIQSPGRNSISPGRNGISPPN 2339
QY 2341 KLSQLPRTSPSTASTSSGSKMSYTSRGRQSQNLTKQTGLSKNASSIPRSESASKG 2400
DB 2340 KLSQLPRTSPSTASTSSGSKMSYTSRGRQSQNLTKQTGLSKNASSIPRSESASKG 2399
QY 2401 LNQMNGNGANKKVELSRMSTSSGSDRSRPPVLVROSTTKEAPSPTLRKLKLEESA 2460
DB 2400 LNQMNGNGANKKVELSRMSTSSGSDRSRPPVLVROSTTKEAPSPTLRKLKLEESA 2459
QY 2461 SFESLSPSPSPASTRQAOQTPVLSPLDMSLSTSSVQAGWRKLPNLSFTIENDG 2520
DB 2460 SFESLSPSPSPASTRQAOQTPVLSPLDMSLSTSSVQAGWRKLPNLSFTIENDG 2519
QY 2521 RPAKRHDIASSHSPSRPLPNRSSTGKREHSHSSSLPRVSTWRRTGSSSSITLSASSES 2580
DB 2520 RPAKRHDIASSHSPSRPLPNRSSTGKREHSHSSSLPRVSTWRRTGSSSSITLSASSES 2579
QY 2581 SEKAKSEDEKXVNSISGTKQSKENQVSAKGTWRKIKENEPSPNTSQTVSSGATNGAES 2640
DB 2580 SEKAKSEDEKXVNSISGTKQSKENQVSAKGTWRKIKENEPSPNTSQTVSSGATNGAES 2639
QY 2641 KTLIYQAPAVSKTDEYVWRIEDCPINNPSGRSPGTGTPPVTDVSEKANPNIKDOKN 2700
DB 2640 KTLIYQAPAVSKTDEYVWRIEDCPINNPSGRSPGTGTPPVTDVSEKANPNIKDOKN 2699
QY 2701 QAKQNVGSGVPMRTVGLNRLNSFIQVADPQKGTIEKPGQNNPVPVSTNETSIVERT 2760
DB 2700 QAKQNVGSGVPMRTVGLNRLNSFIQVADPQKGTIEKPGQNNPVPVSTNETSIVERT 2759
QY 2761 PFSSSSSKHSSPGCTVAARVTPNPNPSRKSADSTSPSCQIPTPVNNNTKEDSKT 2820
DB 2760 PFSSSSSKHSSPGCTVAARVTPNPNPSRKSADSTSPSCQIPTPVNNNTKEDSKT 2819
QY 2821 DSTESSGCTQSPKXHSGLVTSV 2843
DB 2820 DSTESSGCTQSPKXHSGLVTSV 2842

```

RESULT 12

US-07-741-940-2

; Sequence 2, Application US/07741940

; Patent No. 5352775

```

/ GENERAL INFORMATION:
/ APPLICANT: ALBERTSEN, HANS
/ APPLICANT: ANAND, RAKESH
/ APPLICANT: CARLSON, MARY
/ APPLICANT: GRODEN, JOANNA
/ APPLICANT: HEDGE, PHILIP J.
/ APPLICANT: JOSLYN, GEOFF
/ APPLICANT: KINZLER, KENNETH
/ APPLICANT: MARKHAM, ALEXANDER F.
/ APPLICANT: NAKAMURA, YUSUKE
/ APPLICANT: TELIVERIS, ANDREW
/ TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC
/ TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS
/ NUMBER OF SEQUENCES: 94
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Banner, Birch, McKie & Beckett
/ STREET: 1001 G Street, NW
/ CITY: Washington
/ STATE: D.C.
/ COUNTRY: USA
/ ZIP: 20001-4598
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA: US/07/741,940
/ APPLICATION NUMBER: 32,141
/ FILING DATE: 19920109
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Kagan, Sarah A.
/ REGISTRATION NUMBER: 1107.035574
/ REFERENCE/DOCKET NUMBER: 1107.035574
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 202-508-9100
/ TELEFAX: 202-508-9299
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2843 amino acids
/ TYPE: AMINO ACID
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/
US-07-741-940-2

```

```

Query Match 99.8%; Score 14533; DB 1; Length 2843;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2836; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 MAAASYDQLLKQVAKALXNENLNRLQLEDNSNHLTKLETSANMKVILKQLQGSIEDEAM 60
DB 1 MAAASYDQLLKQVAKALXNENLNRLQLEDNSNHLTKLETSANMKVILKQLQGSIEDEAM 60

QY 61 ASSGQIDLLLEKLKLNLDSSNFPKVLRSKMSLSYSGRSGSVSRGECSPVPMGSPFR 120
DB 61 ASSGQIDLLLEKLKLNLDSSNFPKVLRSKMSLSYSGRSGSVSRGECSPVPMGSPFR 120

QY 121 RGFVNGSRESTGYLEELKERSLLADLDEKEKDWYQAQLNLTKRIDSPLTENFSL 180
DB 121 RGFVNGSRESTGYLEELKERSLLADLDEKEKDWYQAQLNLTKRIDSPLTENFSL 180

QY 181 QTDWTRRQLEVEARQIRVAMEEQGTQDMKRAQRIARIQQIEKDLIRQLQSOAT 240
DB 181 QTDWTRRQLEVEARQIRVAMEEQGTQDMKRAQRIARIQQIEKDLIRQLQSOAT 240

QY 241 EAERSSQNKHETGSHDAERQNEGQVGEINMATSGNQGSSTRMDHETASVLSSTHSA 300
DB 241 EAERSSQNKHETGSHDAERQNEGQVGEINMATSGNQGSSTRMDHETASVLSSTHSA 300

QY 301 PRLTSHLGTVMYVYLLSMLGTHDKDMSRTLLAMSSSSQDSCISMRSQSGCLPLLIQLL 360
DB 301 PRLTSHLGTVMYVYLLSMLGTHDKDMSRTLLAMSSSSQDSCISMRSQSGCLPLLIQLL 360

```

QY 361 HGNDKDSVLLGNSRGSKEARARASAAALHNIHSQDDPKRGREIRVTLHLBOIRAYCETC 420  
DB 361 HGNDKDSVLLGNSRGSKEARARASAAALHNIHSQDDPKRGREIRVTLHLBOIRAYCETC 420  
QY 421 WEQBAHEFGMDQKNMPAPVEHQICPAVCLMKLSFDEEHRHAMNELGGIQAIAELLQ 480  
DB 421 WEQBAHEFGMDQKNMPAPVEHQICPAVCLMKLSFDEEHRHAMNELGGIQAIAELLQ 480  
QY 481 VCEMYGLTNDHYSITLREYAGWALTNLTFGVDANKATILCSMKGCWRALVAQLKSESDL 540  
DB 481 VCEMYGLTNDHYSITLREYAGWALTNLTFGVDANKATILCSMKGCWRALVAQLKSESDL 540  
QY 541 QOVIASVLRNLNLSWRADVNSKTLREYVGSVKALMECALEVKKSESTLKSVLNLSALNLSAHT 600  
DB 541 QOVIASVLRNLNLSWRADVNSKTLREYVGSVKALMECALEVKKSESTLKSVLNLSALNLSAHT 600  
QY 601 ENKADICAVDGALEFLVGLTTRYSQNTLAIIESGGILRNVSLLIATNEDHRQILRENN 660  
DB 601 ENKADICAVDGALEFLVGLTTRYSQNTLAIIESGGILRNVSLLIATNEDHRQILRENN 660  
QY 661 CIQTLLOHLKSHSLATVSNACGTLNLSARNPKDOEALWDMGAVSNLXNLIHSKHKMIAM 720  
DB 661 CIQTLLOHLKSHSLATVSNACGTLNLSARNPKDOEALWDMGAVSNLXNLIHSKHKMIAM 720  
QY 721 GSAALRNLMANRPKYKDANINSGSSLPSSHVKQKALEAELDAOHLSETFDNIDLS 780  
DB 721 GSAALRNLMANRPKYKDANINSGSSLPSSHVKQKALEAELDAOHLSETFDNIDLS 780  
QY 781 PKASHRSKORHKOSLYGDFVDTNRHDDNRSNENTGNMTVLSPYLNTTIVLPSSSSSRGS 840  
DB 781 PKASHRSKORHKOSLYGDFVDTNRHDDNRSNENTGNMTVLSPYLNTTIVLPSSSSSRGS 840  
QY 841 LBSRSSEKDRSLERERGI GLGNVHPATENPGTSSKRGILQISTAAQIAKWEVEVSAIHTS 900  
DB 841 LBSRSSEKDRSLERERGI GLGNVHPATENPGTSSKRGILQISTAAQIAKWEVEVSAIHTS 900  
QY 901 QEDRSSGSTTELHCVTDERNALRRSAATHNTYNTFKSENSNRTCSMPYAKLEYKSS 960  
DB 901 QEDRSSGSTTELHCVTDERNALRRSAATHNTYNTFKSENSNRTCSMPYAKLEYKSS 960  
QY 961 NDLNSVSSDGYGKRGQMKPSIESYSEDDSKFCYGOYPADLAHKIHSANHMDNDGE 1020  
DB 961 NDLNSVSSDGYGKRGQMKPSIESYSEDDSKFCYGOYPADLAHKIHSANHMDNDGE 1020  
QY 1021 LPTPINYSIKYSDQNLNSGRQSPQNERWARPKHIIIEBKQSEQRQRNOSTTYPVYTE 1080  
DB 1021 LPTPINYSIKYSDQNLNSGRQSPQNERWARPKHIIIEBKQSEQRQRNOSTTYPVYTE 1080  
QY 1081 STDDKHLKFPQHFQGOECVSPVRSRANGSETNRVGSNHGINQNVQSICQEDDVEDDKP 1140  
DB 1081 STDDKHLKFPQHFQGOECVSPVRSRANGSETNRVGSNHGINQNVQSICQEDDVEDDKP 1140  
QY 1141 TNYSEYSEEEQHEBERPTNYSIKYNEBKRVDPIDYSLKYATDIPSSQKQSFSPKS 1200  
DB 1141 TNYSEYSEEEQHEBERPTNYSIKYNEBKRVDPIDYSLKYATDIPSSQKQSFSPKS 1200  
QY 1201 SSGQSKTEHMSSENSTTPSNAKRQNLHPNSAQSRSGOPOKAATCKVSSINQETIQ 1260  
DB 1201 SSGQSKTEHMSSENSTTPSNAKRQNLHPNSAQSRSGOPOKAATCKVSSINQETIQ 1260  
QY 1261 TYCVEDTPICFRCSLSLSAEDBIGCNQTTQADSAANTLQIAEIKKIGTRAEADPV 1320  
DB 1261 TYCVEDTPICFRCSLSLSAEDBIGCNQTTQADSAANTLQIAEIKKIGTRAEADPV 1320  
QY 1321 SEVPVAVSQHPRTKSSRLOGSSLSSESARHKAVEFPSSGAKSPKSGAQTPKSPPEHYVQET 1380  
DB 1321 SEVPVAVSQHPRTKSSRLOGSSLSSESARHKAVEFPSSGAKSPKSGAQTPKSPPEHYVQET 1380  
QY 1381 PLMFBSCTSVSLDPSFESSIASVCSFPCSGMWGIIISPSLDPSPGQTMPPSSKTPP 1440  
DB 1381 PLMFBSCTSVSLDPSFESSIASVCSFPCSGMWGIIISPSLDPSPGQTMPPSSKTPP 1440  
QY 1441 PPPQTAQTKEVPKXKAPTAEKRESGPKQAANAAVQVLPDADTLLHFATESTPDGP 1500

DB 1441 PPPQTAQTKEVPKXKAPTAEKRESGPKQAANAAVQVLPDADTLLHFATESTPDGP 1500  
QY 1501 SCSSLSALSILDEPFFIQKQVELRIMPVQENDNGNETESQPKESNENQKEAKETIDSE 1560  
DB 1501 SCSSLSALSILDEPFFIQKQVELRIMPVQENDNGNETESQPKESNENQKEAKETIDSE 1560  
QY 1561 KDLDDSDDDDDIETLBECHISAMPTSSRKAKPAQATASKLPPVARKPQOLPVYKLLPS 1620  
DB 1561 KDLDDSDDDDDIETLBECHISAMPTSSRKAKPAQATASKLPPVARKPQOLPVYKLLPS 1620  
QY 1621 QNRLOPKHVSFTPGDDMPRVYCVEGTPIINFSTATSLSDLTIESPPNELAAGVGVRGAQ 1680  
DB 1621 QNRLOPKHVSFTPGDDMPRVYCVEGTPIINFSTATSLSDLTIESPPNELAAGVGVRGAQ 1680  
QY 1681 SGEFEKEDTTPTEGRSTDEAOGGKTSSVTPELDDNKABEGDILAEICINAMPKGSHKP 1740  
DB 1681 SGEFEKEDTTPTEGRSTDEAOGGKTSSVTPELDDNKABEGDILAEICINAMPKGSHKP 1740  
QY 1741 PRVKKIMQVQOASASSAPKNQLOGKKKPKTPSVKPIPQNTTEYRTRVRKNADSKNNLN 1800  
DB 1741 PRVKKIMQVQOASASSAPKNQLOGKKKPKTPSVKPIPQNTTEYRTRVRKNADSKNNLN 1800  
QY 1801 AERVFSNKNKONLKNKSKDFNDKLPNNEDRVRSFAFDSPHHYTPIEGTTPYCFSRND 1860  
DB 1801 AERVFSNKNKONLKNKSKDFNDKLPNNEDRVRSFAFDSPHHYTPIEGTTPYCFSRND 1860  
QY 1861 SLSLDDDDDDVLSREKAEALRKAENKESAEKVTSHTLTSTNQOSANKTOAIKQPINR 1920  
DB 1861 SLSLDDDDDDVLSREKAEALRKAENKESAEKVTSHTLTSTNQOSANKTOAIKQPINR 1920  
QY 1921 GQPKFILQKOSTPPQSSKDIIPDRGAATDEKLOQPAIENTPVCFSHNSLSLSDIDQENN 1980  
DB 1921 GQPKFILQKOSTPPQSSKDIIPDRGAATDEKLOQPAIENTPVCFSHNSLSLSDIDQENN 1980  
QY 1981 NKNEPIKETEPDPSQGEPSKQASGYAPKSFVVEDTTPVCFSNSSLSLSISEDDLLQ 2040  
DB 1981 NKNEPIKETEPDPSQGEPSKQASGYAPKSFVVEDTTPVCFSNSSLSLSISEDDLLQ 2040  
QY 2041 ECISAMPKKKKPSRLKGDNEKHSRNMGGILGEBDLTDLKDQRPDSEHGLSPDSNFD 2100  
DB 2041 ECISAMPKKKKPSRLKGDNEKHSRNMGGILGEBDLTDLKDQRPDSEHGLSPDSNFD 2100  
QY 2101 WKAIQEGANSIVSSLHAAAAACLSROASSDSDSLILSKGILSGSPFHTPDQEEKPT 2160  
DB 2101 WKAIQEGANSIVSSLHAAAAACLSROASSDSDSLILSKGILSGSPFHTPDQEEKPT 2160  
QY 2161 SNKGPRILKPGEXSTLTETKIESESKGKGGKVKYKSLITGKVRNSSEISGQMKQPLQAN 2220  
DB 2161 SNKGPRILKPGEXSTLTETKIESESKGKGGKVKYKSLITGKVRNSSEISGQMKQPLQAN 2220  
QY 2221 MPSISRGRTHIHPGVNNSSTSPVSKGPPKTPASKSPSEGOTATTSPRGAKPSVK 2280  
DB 2221 MPSISRGRTHIHPGVNNSSTSPVSKGPPKTPASKSPSEGOTATTSPRGAKPSVK 2280  
QY 2281 ELSVARSQTCIGSSKAPSRSGRSTPSRPAQOPLSRPIQSPGRNSISPGRNGISPPN 2340  
DB 2281 ELSVARSQTCIGSSKAPSRSGRSTPSRPAQOPLSRPIQSPGRNSISPGRNGISPPN 2340  
QY 2341 KLSQLPRTSPSTASTKSSGSGKMSYTPSGRQMSQQNLTKQTGLSKNASSIPRESASKG 2400  
DB 2341 KLSQLPRTSPSTASTKSSGSGKMSYTPSGRQMSQQNLTKQTGLSKNASSIPRESASKG 2400  
QY 2401 LNQMNGNGANKVELSRMSSTKSSGSESDRSPVLRQSTFIKEAPSTLRKLEESA 2460  
DB 2401 LNQMNGNGANKVELSRMSSTKSSGSESDRSPVLRQSTFIKEAPSTLRKLEESA 2460  
QY 2461 SFESLSPSSRPASPTRQAOTPVLSPLPDMSLSTHSSVQAGGWRKLPPLNLSPTIEYNDG 2520  
DB 2461 SFESLSPSSRPASPTRQAOTPVLSPLPDMSLSTHSSVQAGGWRKLPPLNLSPTIEYNDG 2520  
QY 2521 RPAKXHDIAKSHSPSRRLPINRSGTWKRBHSHKSSLPVRSVWRTRTGSSSSILSASSES 2580



Db 2521 RPAKRHDIAHSHSPRLPINRSGTWKREHSHSSSLPRVSTWRRTGSSSSILSASSPS 2580  
Qy 2581 SEKAHSEDEKHVNSISGTQSKENQVSAKGTWRKIKENEFSPNTSQTSSGATNGAES 2640  
Db 2581 SEKAHSEDEKHVNSISGTQSKENQVSAKGTWRKIKENEFSPNTSQTSSGATNGAES 2640  
Qy 2641 KTLIYQAPAVSKTEDVWVRIEDCPINNPRSGRSPGTNTPPVIDSVSEKAPNPKDSKON 2700  
Db 2641 KTLIYQAPAVSKTEDVWVRIEDCPINNPRSGRSPGTNTPPVIDSVSEKAPNPKDSKON 2700  
Qy 2701 QAKONVNGSVPMRTVGLNRLNSFIOVDAPDOKGTGTEIKPQNNPVPVSETNESSIVERT 2760  
Db 2701 QAKONVNGSVPMRTVGLNRLNSFIOVDAPDOKGTGTEIKPQNNPVPVSETNESSIVERT 2760  
Qy 2761 PFSSSSSKHSSPSGTVAAARVTPFNYPNPSRKSADSTARSPPQIPTPVNNNTKKSDDSKT 2820  
Db 2761 PFSSSSSKHSSPSGTVAAARVTPFNYPNPSRKSADSTARSPPQIPTPVNNNTKKSDDSKT 2820  
Qy 2821 DSTESSGTQSPKHSGLYATSV 2843  
Db 2821 DSTESSGTQSPKHSGLYATSV 2843

RESULT 13  
US-08-289-548A-2  
; Sequence 2, Application US/08289548A  
; Patent No. 5648212  
; GENERAL INFORMATION:  
; APPLICANT: ALBERTSEN, HANS  
; APPLICANT: ANAND, RAKESH  
; APPLICANT: CARLSON, MARY  
; APPLICANT: GEODEN, JOANNA  
; APPLICANT: HEDGE, PHILIP J.  
; APPLICANT: JOSLYN, GEOFF  
; APPLICANT: KINZLER, KENNETH  
; APPLICANT: MARKHAM, ALEXANDER F.  
; APPLICANT: NAKAMURA, YUSUKE  
; APPLICANT: THLIVERIS, ANDREW  
; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC  
; TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS  
; NUMBER OF SEQUENCES: 102  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Banner & Allegretti, LTD  
; STREET: 1001 G Street, NW  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20001-4598  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA: US/08/289,548A  
; FILING DATE: 12-AUG-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kagan, Sarah A.  
; REGISTRATION NUMBER: 32,141  
; REFERENCE/DOCKET NUMBER: 1107.46943  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-508-9100  
; TELEFAX: 202-508-9299  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2843 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-289-548A-2  
Query Match 99.8%; Score 14533; DB 1; Length 2843;

Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2836; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
Qy 1 MAAASYDOLLKQVEALKMNENLNQLEEDNSNHLTKLETEASNKMKVILKQIQSGSIEDEAM 60  
Db 1 MAAASYDOLLKQVEALKMNENLNQLEEDNSNHLTKLETEASNKMKVILKQIQSGSIEDEAM 60  
Qy 61 ASSGQIDLLERLKEINLDSNPPGVKURSKWLSYSGSRGSGSVSRSGECSPVPMGSPFR 120  
Db 61 ASSGQIDLLERLKEINLDSNPPGVKURSKWLSYSGSRGSGSVSRSGECSPVPMGSPFR 120  
Qy 121 RGFVNGSRESSTGYLEELKERSLLLLADLDKEEKDWMYYAOLNLTKEIDSLPTEPNFSL 180  
Db 121 RGFVNGSRESSTGYLEELKERSLLLLADLDKEEKDWMYYAOLNLTKEIDSLPTEPNFSL 180  
Qy 181 QTDLTRRQLEEARQIRVAMEEQLGTCDMEKRAORRIARIQOIEKOTILRIQLQSOAT 240  
Db 181 QTDLTRRQLEEARQIRVAMEEQLGTCDMEKRAORRIARIQOIEKOTILRIQLQSOAT 240  
Qy 241 EAERSSQNKHETGSHDAERONEGQGVGEINMATSNGCGSTTRVDHETASVLSSTHSA 300  
Db 241 EAERSSQNKHETGSHDAERONEGQGVGEINMATSNGCGSTTRVDHETASVLSSTHSA 300  
Qy 301 PRLTSLHGTKEVWVYLLSMLGTHDKDDMSRTLLAMSSSDSCISMRQSGCLPILLQLL 360  
Db 301 PRLTSLHGTKEVWVYLLSMLGTHDKDDMSRTLLAMSSSDSCISMRQSGCLPILLQLL 360  
Qy 361 HGNDKDSVLLGNRSGSKARARASAAHLNIIHSQPDQKRRREIRVLHLEQIRAYCETC 420  
Db 361 HGNDKDSVLLGNRSGSKARARASAAHLNIIHSQPDQKRRREIRVLHLEQIRAYCETC 420  
Qy 421 WEWQEAHEPQMDQDNMPAPVEHQICPAVCVLMKLSFDEEHRHMMELGGLQAIABLLQ 480  
Db 421 WEWQEAHEPQMDQDNMPAPVEHQICPAVCVLMKLSFDEEHRHMMELGGLQAIABLLQ 480  
Qy 481 VDCEMYGLTNDHYSITLRYAGMALTNLTFGDVANKATLCMKMGCMRALVAQLKSEBDL 540  
Db 481 VDCEMYGLTNDHYSITLRYAGMALTNLTFGDVANKATLCMKMGCMRALVAQLKSEBDL 540  
Qy 541 COVITASVLNLSRADVNSKKTLEVGSKALMECALEVKKESTLKSVLNLSNLSAHT 600  
Db 541 COVITASVLNLSRADVNSKKTLEVGSKALMECALEVKKESTLKSVLNLSNLSAHT 600  
Qy 601 ENKADICAVDGAFLVGTLYRSQTNLTALIESGGGILRNVSLLIATNEDHRQILRENN 660  
Db 601 ENKADICAVDGAFLVGTLYRSQTNLTALIESGGGILRNVSLLIATNEDHRQILRENN 660  
Qy 661 CLQTLLOHLKSHSLTIVSNACGTLNLSARNPKDOEALWDMGAVSMKLNLIHSHKMIAM 720  
Db 661 CLQTLLOHLKSHSLTIVSNACGTLNLSARNPKDOEALWDMGAVSMKLNLIHSHKMIAM 720  
Qy 721 GSAALRNLMANRPAYKDANIMSPGSSLPVSHVRKQKALEAEALDAQHLSETFDNIDLS 780  
Db 721 GSAALRNLMANRPAYKDANIMSPGSSLPVSHVRKQKALEAEALDAQHLSETFDNIDLS 780  
Qy 781 PKASHRSQRHKSQSLYGDYVDFTRHDDNRSDNFTGNMTVLSPLYLNTTVLPSSSSRGS 840  
Db 781 PKASHRSQRHKSQSLYGDYVDFTRHDDNRSDNFTGNMTVLSPLYLNTTVLPSSSSRGS 840  
Qy 841 LDSRSEKDRSLERERGI GLGNYPATENPOTSKRGLOISTTAAQAKWMEVSAIHTS 900  
Db 841 LDSRSEKDRSLERERGI GLGNYPATENPOTSKRGLOISTTAAQAKWMEVSAIHTS 900  
Qy 901 QEDRSSGTTTELHCVTDERNALRSSAAHTHSNTYNTFKSENSNRTCSMPYAKLEYKRSS 960  
Db 901 QEDRSSGTTTELHCVTDERNALRSSAAHTHSNTYNTFKSENSNRTCSMPYAKLEYKRSS 960  
Qy 961 NDSLNSVSSDGYKRGOMKPSIESYFDDDESFCISYGOVPADLAHAKHSHANHDDNDGE 1020  
Db 961 NDSLNSVSSDGYKRGOMKPSIESYFDDDESFCISYGOVPADLAHAKHSHANHDDNDGE 1020  
Qy 1021 LDTPINYSKYSDQOLNSGRQSPSQNERWAPKHIIEDEIKQSPQRQSRNSTTYPVYTE 1080



Db 1021 LDPTINYLKYSDEGLNSGROSPSONEMARPKHIEDEIKQSORGRNOSTTYPYVTE 1080  
 QY 1081 STDDHKLKFOHPGQOECVSPYRSRGANGSEPTNRVGNHGINOVSOGLCOEDDYEDDKP 1140  
 Db 1081 STDDHKLKFOHPGQOECVSPYRSRGANGSEPTNRVGNHGINOVSOGLCOEDDYEDDKP 1140  
 QY 1141 TNSRYRSEEOHHEERPTVYSTKYNEEKRHVOPIDYSLKYATDPIPSGOKSPFSKS 1200  
 Db 1141 TNSRYRSEEOHHEERPTVYSTKYNEEKRHVOPIDYSLKYATDPIPSGOKSPFSKS 1200  
 QY 1201 SSGQSKTEHMSSESENTSTPSNAKRONQLPSSAQSRGQPKAATCKXSSINQETIQ 1260  
 Db 1201 SSGQSKTEHMSSESENTSTPSNAKRONQLPSSAQSRGQPKAATCKXSSINQETIQ 1260  
 QY 1261 TYCEDTPICFRSCGSLSSABDEICNOTTOBADSANTLOJAEIKXGTSAEDPV 1320  
 Db 1261 TYCEDTPICFRSCGSLSSABDEICNOTTOBADSANTLOJAEIKXGTSAEDPV 1320  
 QY 1321 SEVPAVSQHPRTKSSRLQSSLSSESAHKAVERSSGAKSPKSGAQTPKSPZHYVQET 1380  
 Db 1321 SEVPAVSQHPRTKSSRLQSSLSSESAHKAVERSSGAKSPKSGAQTPKSPZHYVQET 1380  
 QY 1381 PLMSRCTSVSLDSFESRSIASSVQSPGCMVSGTISPDLPDSPOQTWPPERSKTPP 1440  
 Db 1381 PLMSRCTSVSLDSFESRSIASSVQSPGCMVSGTISPDLPDSPOQTWPPERSKTPP 1440  
 QY 1441 PPQTOJOTKREVPKXKAFTAEKRESGPKOAVNAVORVOYLPAADTLHFAETESTPDGF 1500  
 Db 1441 PPQTOJOTKREVPKXKAFTAEKRESGPKOAVNAVORVOYLPAADTLHFAETESTPDGF 1500  
 QY 1501 SCSSLSLALSDPEFIQCVBLRIMPVQENDNGNETESBQPKESNENOEKAEKTIJSE 1560  
 Db 1501 SCSSLSLALSDPEFIQCVBLRIMPVQENDNGNETESBQPKESNENOEKAEKTIJSE 1560  
 QY 1561 KDLDSDSDDDIIELEBCTIISAMPTKSRKAKPAQATSKLPPVAKPSPQLPYKLLPS 1620  
 Db 1561 KDLDSDSDDDIIELEBCTIISAMPTKSRKAKPAQATSKLPPVAKPSPQLPYKLLPS 1620  
 QY 1621 QNRLOPQKHVSFTTGDDKPRVYCVGEPINFTATSLSDLTIESPPMLAAGEVIRGAQ 1680  
 Db 1621 QNRLOPQKHVSFTTGDDKPRVYCVGEPINFTATSLSDLTIESPPMLAAGEVIRGAQ 1680  
 QY 1681 SGEFERKDTITTEBGRSTDEAOGKTSVTLPELDONKAEBSGDIABECINSAMPKSHKP 1740  
 Db 1681 SGEFERKDTITTEBGRSTDEAOGKTSVTLPELDONKAEBSGDIABECINSAMPKSHKP 1740  
 QY 1741 FRVKKINDVOOASASSAPKNOJDKKKKPTSPVXIIPONTETRTVRKNASKNLN 1800  
 Db 1741 FRVKKINDVOOASASSAPKNOJDKKKKPTSPVXIIPONTETRTVRKNASKNLN 1800  
 QY 1801 AERYFSNDKDSKONLXNKSXDFNDKLPNNEDRVKGSFAPDSPHYTPIEGTPLYCFGRND 1860  
 Db 1801 AERYFSNDKDSKONLXNKSXDFNDKLPNNEDRVKGSFAPDSPHYTPIEGTPLYCFGRND 1860  
 QY 1861 SLSSLDPDDDVDLSREKAEIRKAKENKESAKTSHETLSNOOSANKTOAIKOPINR 1920  
 Db 1861 SLSSLDPDDDVDLSREKAEIRKAKENKESAKTSHETLSNOOSANKTOAIKOPINR 1920  
 QY 1921 GQPKPILOKOSTFPOSSKDIIDRGATDEKONFAIENTPVCFSHNSLSLSIDEDILQ 1980  
 Db 1921 GQPKPILOKOSTFPOSSKDIIDRGATDEKONFAIENTPVCFSHNSLSLSIDEDILQ 1980  
 QY 1981 NKEBPIETETPPPOGEPKPOASGYAPKSFHVEDTPVCFSRNSSLSSIDEDILQ 2040  
 Db 1981 NKEBPIETETPPPOGEPKPOASGYAPKSFHVEDTPVCFSRNSSLSSIDEDILQ 2040  
 QY 2041 ECISGAMPKKKKPSRLKGDNEKSPRNMGILGEDLTLDKDIORPSEHGLSDSENFD 2100  
 Db 2041 ECISGAMPKKKKPSRLKGDNEKSPRNMGILGEDLTLDKDIORPSEHGLSDSENFD 2100  
 QY 2101 WKAIOEGANSIVSSIHQAAAAACISROASSDSDSLSLKSGISLGSPPHLLTPDOEKXPFT 2160  
 Db 2101 WKAIOEGANSIVSSIHQAAAAACISROASSDSDSLSLKSGISLGSPPHLLTPDOEKXPFT 2160

QY 2161 SNKGPRILKPEKSTLETYKIESBKGIGKAKVYKSLITGKVRNSBIISGOMKOPLOAN 2220  
 Db 2161 SNKGPRILKPEKSTLETYKIESBKGIGKAKVYKSLITGKVRNSBIISGOMKOPLOAN 2220  
 QY 2221 MPEISRGTMHHPGVNRNSSSTSPVSKKGPPLTKPASKSPSEGOATISPGKAXSVKS 2280  
 Db 2221 MPEISRGTMHHPGVNRNSSSTSPVSKKGPPLTKPASKSPSEGOATISPGKAXSVKS 2280  
 QY 2281 ELSFVARQTSQIGSSKADSRGSGRDSRSPRPAQOPLASPIOSPGRNSISPGRNGISPN 2340  
 Db 2281 ELSFVARQTSQIGSSKADSRGSGRDSRSPRPAQOPLASPIOSPGRNSISPGRNGISPN 2340  
 QY 2341 KLSQLPRTSSPTASTKSSGSGKMTSPGRQMSQONLTQOTGLSKMASIPRBSBASKG 2400  
 Db 2341 KLSQLPRTSSPTASTKSSGSGKMTSPGRQMSQONLTQOTGLSKMASIPRBSBASKG 2400  
 QY 2401 INOMNNGANKKVEJSRMSTKSSGSESDRSERPVLVQSTFIKEAPSPTLRKLBEBA 2460  
 Db 2401 INOMNNGANKKVEJSRMSTKSSGSESDRSERPVLVQSTFIKEAPSPTLRKLBEBA 2460  
 QY 2461 SFESLSPPSRPASPTRSQAQTPVLSPLSDMLSTHSSVQAGGMRKLPNLSPTTEYNDG 2520  
 Db 2461 SFESLSPPSRPASPTRSQAQTPVLSPLSDMLSTHSSVQAGGMRKLPNLSPTTEYNDG 2520  
 QY 2521 RPAKRHDIDARSSESPSRPLPNRSGTWKREHSHSSLPVSVGMRRTGSSSSTLSASSBS 2580  
 Db 2521 RPAKRHDIDARSSESPSRPLPNRSGTWKREHSHSSLPVSVGMRRTGSSSSTLSASSBS 2580  
 QY 2581 SEKAKSEDEKHNVSISGTQKSENOVSAGKTWRKIKENEFSPNTSQTQVSSGATNGAS 2640  
 Db 2581 SEKAKSEDEKHNVSISGTQKSENOVSAGKTWRKIKENEFSPNTSQTQVSSGATNGAS 2640  
 QY 2641 KTLIYQMAPAVASTEDVWVRIEDCPINNRSRSGSPGNTPPVIDSSEKAMPNIKSKN 2700  
 Db 2641 KTLIYQMAPAVASTEDVWVRIEDCPINNRSRSGSPGNTPPVIDSSEKAMPNIKSKN 2700  
 QY 2701 QAKQNGVGSVMPRTYGLNRLNLSFIQVADAPQKTEIRPQONNPVAVSETNESSIVER 2760  
 Db 2701 QAKQNGVGSVMPRTYGLNRLNLSFIQVADAPQKTEIRPQONNPVAVSETNESSIVER 2760  
 QY 2761 PRESSSSSHSSPSGVAARVTPFNYPNBPBKSSADTSABSOQPTPVNNTTKRDSKT 2820  
 Db 2761 PRESSSSSHSSPSGVAARVTPFNYPNBPBKSSADTSABSOQPTPVNNTTKRDSKT 2820  
 QY 2821 DSTESSGTQSPKXHSGLVLTSTV 2843  
 Db 2821 DSTESSGTQSPKXHSGLVLTSTV 2843  
 RESULT 14  
 US-08-452-654-2  
 ; Sequence 2, Application US/08452654  
 ; Patent No. 5691454  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ALBERTSEN, HANS  
 ; APPLICANT: ANAND, RAKESH  
 ; APPLICANT: CARLSON, MARY  
 ; APPLICANT: GRODEN, JOANNA  
 ; APPLICANT: HEDGE, PHILIP J.  
 ; APPLICANT: JOSLYN, GEOFF  
 ; APPLICANT: KINZLER, KENNETH  
 ; APPLICANT: MARKAM, ALEXANDER F.  
 ; APPLICANT: NAKAMURA, YUSUKE  
 ; APPLICANT: THILVERIS, ANDREW  
 ; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC  
 ; NUMBER OF SEQUENCES: 94  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Hammer, Birch, McKie & Beckett  
 ; STREET: 1001 G Street, NW  
 ; CITY: Washington  
 ; STATE: D.C.

COUNTRY: USA  
ZIP: 20001-4598  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/452,654  
FILING DATE: 25-MAY-1995  
CLASSIFICATION: 536  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 07/741,940  
FILING DATE: 08-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Kagan, Sarah A.  
REGISTRATION NUMBER: 32,141  
REFERENCE/DOCKET NUMBER: 1107.035574  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-508-9100  
TELEFAX: 202-508-9299  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2843 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-452-654-2

Query Match 99.8%; Score 14533; DB 1; Length 2843;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2836; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 MAAASVDQLKQVEALKEKNSNLTROLEEDNSNHLTKLEFASNMKEVILKOLQGSIEDAM 60  
DB 1 MAAASVDQLKQVEALKEKNSNLTROLEEDNSNHLTKLEFASNMKEVILKOLQGSIEDAM 60  
QY 61 ASSGQIDLLERLKEINLDSNFPGVKRSKMSLRSGSRSRSGSCSPVPWGSFPR 120  
DB 61 ASSGQIDLLERLKEINLDSNFPGVKRSKMSLRSGSRSRSGSCSPVPWGSFPR 120  
QY 121 RGFVNSRSTGTLELEKERSLLADLDKEKEDWYAAQQLTKRIDSLPLENSSL 180  
DB 121 RGFVNSRSTGTLELEKERSLLADLDKEKEDWYAAQQLTKRIDSLPLENSSL 180  
QY 181 QTDNTRQLEVEARQIRVAMEZOLGTCDMEKRAQRIARIQIEXDILIRIOLLOSGAT 240  
DB 181 QTDNTRQLEVEARQIRVAMEZOLGTCDMEKRAQRIARIQIEXDILIRIOLLOSGAT 240  
QY 241 EAEBSQNKERTGSHDAERQNGGVGEINMATSNGGOSTTRMDHETASVLSSTHSA 300  
DB 241 EAEBSQNKERTGSHDAERQNGGVGEINMATSNGGOSTTRMDHETASVLSSTHSA 300  
QY 301 PRLTSHLGTKEYMYSLISM,GTGDKDMSTRLLAMSSODSCISMROSGCLPLLIQL 360  
DB 301 PRLTSHLGTKEYMYSLISM,GTGDKDMSTRLLAMSSODSCISMROSGCLPLLIQL 360  
QY 361 HNDYDVSILGNSRSKKEAPASALAHNIHSQPDYKGRREIFVHLLEQIRAYCETC 420  
DB 361 HNDYDVSILGNSRSKKEAPASALAHNIHSQPDYKGRREIFVHLLEQIRAYCETC 420  
QY 421 MEMOAHBEHGMODKNPAPAVEHQICPAVCYMLSPDEHRHAMNELLGLOAIETLQ 480  
DB 421 MEMOAHBEHGMODKNPAPAVEHQICPAVCYMLSPDEHRHAMNELLGLOAIETLQ 480  
QY 481 VDCENYGLTNDHYSITLRRYAGMALTNLTFGDVANKATLCNMKGCMRALVAOLKSESD 540  
DB 481 VDCENYGLTNDHYSITLRRYAGMALTNLTFGDVANKATLCNMKGCMRALVAOLKSESD 540  
QY 541 QOVIVASVLRNLSPRADVNSKTLBRVGSKALMEGALYKKESTLKYLSALMNSACT 600  
DB 541 QOVIVASVLRNLSPRADVNSKTLBRVGSKALMEGALYKKESTLKYLSALMNSACT 600

QY 601 ENKADIICAVDGAALFVGLTYRSQNTLAIIESGGGILIRNYSILATKEDHROILRENN 660  
DB 601 ENKADIICAVDGAALFVGLTYRSQNTLAIIESGGGILIRNYSILATKEDHROILRENN 660  
QY 661 CLQTLLOHLKSHSLITVSNACGLTMMI,SAKNPDCEALMDGAVSNLXNLIHSKXKIAM 720  
DB 661 CLQTLLOHLKSHSLITVSNACGLTMMI,SAKNPDCEALMDGAVSNLXNLIHSKXKIAM 720  
QY 721 GSAALRLNMANPACTYKANTMSPSSLPULHYRQKLELDAQHLSSEFNDINDLS 780  
DB 721 GSAALRLNMANPACTYKANTMSPSSLPULHYRQKLELDAQHLSSEFNDINDLS 780  
QY 781 PKASHRSKQKHQSLGYVFDINRADDNRSDNFTGNMTVASPYLNTLVPSSSSRGS 840  
DB 781 PKASHRSKQKHQSLGYVFDINRADDNRSDNFTGNMTVASPYLNTLVPSSSSRGS 840  
QY 841 LDSRSSEKORSLEREBICIGLGNYPATENPGISKRGLQISTTAQIACVMEVSLIHS 900  
DB 841 LDSRSSEKORSLEREBICIGLGNYPATENPGISKRGLQISTTAQIACVMEVSLIHS 900  
QY 901 QEDRSSGSTTELCVDERNALRRSSAATHSTNTYFTSENSNRCTSPYAKLEYKRS 960  
DB 901 QEDRSSGSTTELCVDERNALRRSSAATHSTNTYFTSENSNRCTSPYAKLEYKRS 960  
QY 961 NDSLSNVSSSDGYGKRGQMKPSTIESYEDDESKFCYGYPADLAHKTISHANMDNCE 1020  
DB 961 NDSLSNVSSSDGYGKRGQMKPSTIESYEDDESKFCYGYPADLAHKTISHANMDNCE 1020  
QY 1021 LDPPIVSLKYSDQOLNSGRQSPQNERMARPKHIIEDDIKQSEBORQSNQSTTPYVTE 1080  
DB 1021 LDPPIVSLKYSDQOLNSGRQSPQNERMARPKHIIEDDIKQSEBORQSNQSTTPYVTE 1080  
QY 1081 STDQKHLKQPHFGQOEVSYPYRSRGANGSEINRVGNSNGINQVNSQSLCOEDDYEDKP 1140  
DB 1081 STDQKHLKQPHFGQOEVSYPYRSRGANGSEINRVGNSNGINQVNSQSLCOEDDYEDKP 1140  
QY 1141 TNYSERYSEREEQEEERPTNHSIKNEKRAHYDQIDYSLKANDIPSSQVQSFSGS 1200  
DB 1141 TNYSERYSEREEQEEERPTNHSIKNEKRAHYDQIDYSLKANDIPSSQVQSFSGS 1200  
QY 1201 SSGQSKTEHMSSSSNTSTPSSNAKRONQLHPSASQSGOPKATCKVNSINQETIO 1260  
DB 1201 SSGQSKTEHMSSSSNTSTPSSNAKRONQLHPSASQSGOPKATCKVNSINQETIO 1260  
QY 1261 TYCEVEDTPICFRSCSLSSLSAEDBIGCQNTTOEADSANTLOIAIKKIGTRSAEDV 1320  
DB 1261 TYCEVEDTPICFRSCSLSSLSAEDBIGCQNTTOEADSANTLOIAIKKIGTRSAEDV 1320  
QY 1321 SEVPAYSOHPRTKSRLOQSSLSSEBSARKHAYEFSGASPKSGAQTKSPPEHYQET 1380  
DB 1321 SEVPAYSOHPRTKSRLOQSSLSSEBSARKHAYEFSGASPKSGAQTKSPPEHYQET 1380  
QY 1381 PLMFSRCTSVSLDSFESSRSLASSVQSEPCGMSGIIISPSDLPDSPGQTMPPSRKTPP 1440  
DB 1381 PLMFSRCTSVSLDSFESSRSLASSVQSEPCGMSGIIISPSDLPDSPGQTMPPSRKTPP 1440  
QY 1441 PPQOTQKREVKNAFPAERKESGPKQAAVAAVQVQVLPDADTLHPFTESIPDGF 1500  
DB 1441 PPQOTQKREVKNAFPAERKESGPKQAAVAAVQVQVLPDADTLHPFTESIPDGF 1500  
QY 1501 SCSSSLASLDEPFIQKQVEALRIMPVQENDNGETSEOPKESNENKEAEKTIQSE 1560  
DB 1501 SCSSSLASLDEPFIQKQVEALRIMPVQENDNGETSEOPKESNENKEAEKTIQSE 1560  
QY 1561 KDLLDSDDDDIIELECIISAMPTSSSRKAKKPAOTAKLPPVARKSQOLPYVTLPS 1620  
DB 1561 KDLLDSDDDDIIELECIISAMPTSSSRKAKKPAOTAKLPPVARKSQOLPYVTLPS 1620  
QY 1621 QNRLOPKHVSFTPGDMFRVYCVBETPINFGTATSLDLTISPPNELAAGEGVAGAQ 1680  
DB 1621 QNRLOPKHVSFTPGDMFRVYCVBETPINFGTATSLDLTISPPNELAAGEGVAGAQ 1680  
QY 1681 SGEFERKDTITPYEGRSTDEAQGGKTSVTIIPBLDNKABEGDILAEICINAPKKGSHKP 1740

1681 SCEPEKDTITTEGRSTDEAGCGTTSATVTELDNDNKAEBDILAEINAMPKGKHNP 1740  
1741 FRVKKINDQVOOASASSAPNNKOLDGKKKPTSPVKPIPONTYFRVRKANADSKNNLN 1800  
1741 FRVKKINDQVOOASASSAPNNKOLDGKKKPTSPVKPIPONTYFRVRKANADSKNNLN 1800  
1801 AERYVSDNKGSKONLKANSKDFNDKLPNNEDVRGSAFSPHHTPIPTGTCFGRND 1860  
1801 AERYVSDNKGSKONLKANSKDFNDKLPNNEDVRGSAFSPHHTPIPTGTCFGRND 1860  
1861 SLSSLDPDDDDVDLSREKALRKAKENKESBAKTSHTLETNSQOANKTOALAKOPINR 1920  
1861 SLSSLDPDDDDVDLSREKALRKAKENKESBAKTSHTLETNSQOANKTOALAKOPINR 1920  
1921 GQKPIIQQKSTFPOSSKDIIDRGATDEKIQNTAIENTPVCFSHNSLSLSDIDENN 1980  
1921 GQKPIIQQKSTFPOSSKDIIDRGATDEKIQNTAIENTPVCFSHNSLSLSDIDENN 1980  
1981 NKENEPIKETEPPOSGEPSPKQASGYAPKSFHYEDTPVCFGRNSLSLSDIDEDLLQ 2040  
1981 NKENEPIKETEPPOSGEPSPKQASGYAPKSFHYEDTPVCFGRNSLSLSDIDEDLLQ 2040  
2041 ECISSNAPKKKKSRLLKGDNEKSPRNWGGILGSDLTLDLKDIOGRPSEHGLSPDSNFD 2100  
2041 ECISSNAPKKKKSRLLKGDNEKSPRNWGGILGSDLTLDLKDIOGRPSEHGLSPDSNFD 2100  
2101 WKALOEGANSIVSSILHQAACLRASDSDSILSLKSGISLSPFHLTPDOEKXPT 2160  
2101 WKALOEGANSIVSSILHQAACLRASDSDSILSLKSGISLSPFHLTPDOEKXPT 2160  
2161 SNKSPRIILKPGKSTLETKLIESESKGKGGKXVYKSLITGVANSNEISGOMKQLOAN 2220  
2161 SNKSPRIILKPGKSTLETKLIESESKGKGGKXVYKSLITGVANSNEISGOMKQLOAN 2220  
2221 MPSTSRGTMTHIGVNRSSSTSPVSKGKBPPLTPASKSPSECOCTTSPRAGKBPVKS 2280  
2221 MPSTSRGTMTHIGVNRSSSTSPVSKGKBPPLTPASKSPSECOCTTSPRAGKBPVKS 2280  
2281 ELSPVARQTSQIGSSSKAPSRSGRSDTSPSRPAQPLSRPIQSPGRMSISPRGAGISPPN 2340  
2281 ELSPVARQTSQIGSSSKAPSRSGRSDTSPSRPAQPLSRPIQSPGRMSISPRGAGISPPN 2340  
2341 KLSQLPRTSSPSTASTKSSGSKMSYSPRGOMQONLTQKTGSKVANSIIPRESASKG 2400  
2341 KLSQLPRTSSPSTASTKSSGSKMSYSPRGOMQONLTQKTGSKVANSIIPRESASKG 2400  
2401 LNOMNNGGAKKATYELSPMSSTKSSGSEDSERPVAVROSTFKEAPSPFLRKLESA 2460  
2401 LNOMNNGGAKKATYELSPMSSTKSSGSEDSERPVAVROSTFKEAPSPFLRKLESA 2460  
2461 SFESLSPPSPASPTPSQAQTPVLSPLPDMSLSTHSSVQAGMRKLPENISPTIEVNDG 2520  
2461 SFESLSPPSPASPTPSQAQTPVLSPLPDMSLSTHSSVQAGMRKLPENISPTIEVNDG 2520  
2521 RPARKHDIARSHSPSRPLPINRSGTMRKSHKSSSLPRVSTWRKRGSSSSILASSES 2580  
2521 RPARKHDIARSHSPSRPLPINRSGTMRKSHKSSSLPRVSTWRKRGSSSSILASSES 2580  
2581 SEKASSEDEKAVNISJTKQKENOVSAKGTWRIKENEFSPINSQGVSSGATNGAES 2640  
2581 SEKASSEDEKAVNISJTKQKENOVSAKGTWRIKENEFSPINSQGVSSGATNGAES 2640  
2641 KTLIYQMAPAYSKTEDVWRIEDCPINNPRSGRSPGTGTPVIDSVSEKAPNKIKSDKN 2700  
2641 KTLIYQMAPAYSKTEDVWRIEDCPINNPRSGRSPGTGTPVIDSVSEKAPNKIKSDKN 2700  
2701 QAKQVNGSVPMRTVGLLENLNSFIQVADAPDOKGTEIKPGONNPVPVSETNESIIERT 2760  
2701 QAKQVNGSVPMRTVGLLENLNSFIQVADAPDOKGTEIKPGONNPVPVSETNESIIERT 2760  
2761 PPSSSSSSKHSGSPGTVAAVTPFNYNPSPKSSADSTISAPSQIPTPVNNNTKKRDSKT 2820  
2761 PPSSSSSSKHSGSPGTVAAVTPFNYNPSPKSSADSTISAPSQIPTPVNNNTKKRDSKT 2820

2761 PPSSSSSSKHSGSPGTVAAVTPFNYNPSPKSSADSTISAPSQIPTPVNNNTKKRDSKT 2820  
2821 DSTESSGTOSPKRHSGSYLVTSV 2843  
2821 DSTESSGTOSPKRHSGSYLVTSV 2843  
RESULT 15  
US-08-370-235A-2  
Sequence 2, Application US/08370235A  
Patent No. 5910418  
GENERAL INFORMATION:  
APPLICANT: VOGELSTEIN, BERT  
APPLICANT: KINZLER, KENNETH W.  
APPLICANT: HILL, DAVID E.  
APPLICANT: JOHNSON, KAREN A.  
TITLE OF INVENTION: ANTIBODIES AND ASSAYS FOR DETERMINING  
TITLE OF INVENTION: MUTATIONS IN THE APC GENE  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: BANNER & WITCOFF, LTD.  
STREET: 1001 G STREET, N.W.  
CITY: WASHINGTON  
STATE: DC  
COUNTRY: US  
ZIP: 20001  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
FILING DATE: 01-JAN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: KAGAN, SARAH A.  
REGISTRATION NUMBER: 32,141  
REFERENCE/DOCKET NUMBER: 0107, 48668  
TELEPHONE: 202 508 9299  
TELEFAX: 202 508 9299  
INFORMATION FOR SEQ. ID NO. 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2843 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-370-235A-2  
Query Match 99.8%; Score 14533; DB 2; Length 2843;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2836; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
1 MAASVDDLLKQVEALKMSNSNLRQLEEDNSNHLTCLTEASNMKEVLKQLOGSIIDEAM 60  
1 MAASVDDLLKQVEALKMSNSNLRQLEEDNSNHLTCLTEASNMKEVLKQLOGSIIDEAM 60  
61 ASSGQIDLLERLKEINLDDSNPPGVKLRKMSLRISYSGSEGSVSRSGECSVPVWGSFPR 120  
61 ASSGQIDLLERLKEINLDDSNPPGVKLRKMSLRISYSGSEGSVSRSGECSVPVWGSFPR 120  
121 RGFVNSRSTGTGLTELEKERSLLADLDKPEKEXWYAAQONLTKRIDSPLTENFSL 180  
121 RGFVNSRSTGTGLTELEKERSLLADLDKPEKEXWYAAQONLTKRIDSPLTENFSL 180  
181 QTDMTROLEYEAKRQIVAMEEOLGTQDMERAKORRIARIQIEKDILIRQLQSQAT 240  
181 QTDMTROLEYEAKRQIVAMEEOLGTQDMERAKORRIARIQIEKDILIRQLQSQAT 240  
181 QTDMTROLEYEAKRQIVAMEEOLGTQDMERAKORRIARIQIEKDILIRQLQSQAT 240  
241 EARRSSONHFGSHAEONONGGVGEINMATSGNGCGSTTRMDETAASYSSSTGSA 300  
241 EARRSSONHFGSHAEONONGGVGEINMATSGNGCGSTTRMDETAASYSSSTGSA 300  
241 EARRSSONHFGSHAEONONGGVGEINMATSGNGCGSTTRMDETAASYSSSTGSA 300

QY 301 PRLTSHLGTKEWYIILSMGLTHDKODMSHTLLAMSSQDSICISMRQSGCLPILLIQL 360  
 Db 301 PRLTSHLGTKEWYIILSMGLTHDKODMSHTLLAMSSQDSICISMRQSGCLPILLIQL 360  
 QY 361 HENDSDVTLGNSRSGKEARASALNHIHSQDDDRGREIRIVLHLOIAAYCETC 420  
 Db 361 HENDSDVTLGNSRSGKEARASALNHIHSQDDDRGREIRIVLHLOIAAYCETC 420  
 QY 421 MEMOEAEHSGMDQDXNPMAPVEHQICPAVCLMGLSPEDEHRAMNELGLOIAIELLO 480  
 Db 421 MEMOEAEHSGMDQDXNPMAPVEHQICPAVCLMGLSPEDEHRAMNELGLOIAIELLO 480  
 QY 481 VDCENYGLTNDHYSTTLRRYAGMALTNLTFQDVANKATLCISMGKRMALVQKSESBDL 540  
 Db 481 VDCENYGLTNDHYSTTLRRYAGMALTNLTFQDVANKATLCISMGKRMALVQKSESBDL 540  
 QY 541 QCVIASVJNLISWRADVNSKTLREVGSYKALMECALEVKKESTLKSVALMNLISACT 600  
 Db 541 QCVIASVJNLISWRADVNSKTLREVGSYKALMECALEVKKESTLKSVALMNLISACT 600  
 QY 601 ENKADI CAVDGALATLVGTLTYRSQNTLAIIESGGGILRVVSSILANEDHROILRENN 660  
 Db 601 ENKADI CAVDGALATLVGTLTYRSQNTLAIIESGGGILRVVSSILANEDHROILRENN 660  
 QY 661 CLOTLLOHLKSHSLTIVSNAAGTLMNLSARNPKDOEALMDQVAGMLNLIHSXKMTAM 720  
 Db 661 CLOTLLOHLKSHSLTIVSNAAGTLMNLSARNPKDOEALMDQVAGMLNLIHSXKMTAM 720  
 QY 721 GSAALALNIMANRPAYKODANTMSFGSSLPILHVRKQALAEALDQHLSETPNDLNS 780  
 Db 721 GSAALALNIMANRPAYKODANTMSFGSSLPILHVRKQALAEALDQHLSETPNDLNS 780  
 QY 781 PRAHSRSGORHKOSLYGDPVPTNRHDDNRSDNFTGMNTVLSPLYLNTTVLPSSSSSGS 840  
 Db 781 PRAHSRSGORHKOSLYGDPVPTNRHDDNRSDNFTGMNTVLSPLYLNTTVLPSSSSSGS 840  
 QY 841 LPSRSEKDRSLERERIGLIGNYHPATENPGTSSRGLQISTTAQIAKMEVSAIHTS 900  
 Db 841 LPSRSEKDRSLERERIGLIGNYHPATENPGTSSRGLQISTTAQIAKMEVSAIHTS 900  
 QY 901 QEDRSSGSTTEHACTTDERNALRSSAAHTSNTNFTKSEKSNFTCMFPAKLEYKSS 960  
 Db 901 QEDRSSGSTTEHACTTDERNALRSSAAHTSNTNFTKSEKSNFTCMFPAKLEYKSS 960  
 QY 961 NDLSNSVSSSDGYGRGQMKPISIESYSEDDSKFCSYGOYPAULAHKHSANEMDNDGE 1020  
 Db 961 NDLSNSVSSSDGYGRGQMKPISIESYSEDDSKFCSYGOYPAULAHKHSANEMDNDGE 1020  
 QY 1021 LDTPIINYSIKTSDEQLNSGRQSPSONERWAPKHIIEDEIKQSEGRQGRNSTIYPYTTE 1080  
 Db 1021 LDTPIINYSIKTSDEQLNSGRQSPSONERWAPKHIIEDEIKQSEGRQGRNSTIYPYTTE 1080  
 QY 1081 STDDHLKFOHPFGQECVSPYRSRANGSEFNRGSHNGINQVSSQLCEDDYEDDKP 1140  
 Db 1081 STDDHLKFOHPFGQECVSPYRSRANGSEFNRGSHNGINQVSSQLCEDDYEDDKP 1140  
 QY 1141 TNYSERYSBEEQHEEERPTNYSIKYNEKRAHVDPIDYSLKYATDIPSSQKSFSPKS 1200  
 Db 1141 TNYSERYSBEEQHEEERPTNYSIKYNEKRAHVDPIDYSLKYATDIPSSQKSFSPKS 1200  
 QY 1201 SSGGSKTEHMSSESTSTPSNAKRONOLHPSAOSRSGOPQAACTCKVSSINQETIQ 1260  
 Db 1201 SSGGSKTEHMSSESTSTPSNAKRONOLHPSAOSRSGOPQAACTCKVSSINQETIQ 1260  
 QY 1261 TYCVADTPICFGRCSLSLSLGADEIDICNQTQADASANTLQIAEIKGKITGHSABDPV 1320  
 Db 1261 TYCVADTPICFGRCSLSLSLGADEIDICNQTQADASANTLQIAEIKGKITGHSABDPV 1320  
 QY 1321 SEVPAVSQHPRTKSSRLQGSLSLSSASAHKAVESGAKSPKSAQOTPKSPPEHYOET 1380  
 Db 1321 SEVPAVSQHPRTKSSRLQGSLSLSSASAHKAVESGAKSPKSAQOTPKSPPEHYOET 1380

QY 1381 PLMPFRCTSVSLDSTESRSIASVQSEPCSGWVSGIIEPSPULPDSFGOTMPSPSKTTP 1440  
 Db 1381 PLMPFRCTSVSLDSTESRSIASVQSEPCSGWVSGIIEPSPULPDSFGOTMPSPSKTTP 1440  
 QY 1441 PPOQTQKREVPKXNAPFAEKRESSEGPQOAVNAQVQVLPDADTLHFAETSPDGF 1500  
 Db 1441 PPOQTQKREVPKXNAPFAEKRESSEGPQOAVNAQVQVLPDADTLHFAETSPDGF 1500  
 QY 1501 SCSSLSLALSDPEPTQXVEILRIMPVOENDNGNETSEOPKESNENOEKAETIDSE 1560  
 Db 1501 SCSSLSLALSDPEPTQXVEILRIMPVOENDNGNETSEOPKESNENOEKAETIDSE 1560  
 QY 1561 KDLDDSDDDDIIELECTISAMPKTSRKAAPKAPQATKLPVPYARKSQULPVYKLPBS 1620  
 Db 1561 KDLDDSDDDDIIELECTISAMPKTSRKAAPKAPQATKLPVPYARKSQULPVYKLPBS 1620  
 QY 1621 QNRLQPKHVSFTPGDMMRVYCEGTPIINFSTATSLSDLTIESPPNELAAGEVGAQ 1680  
 Db 1621 QNRLQPKHVSFTPGDMMRVYCEGTPIINFSTATSLSDLTIESPPNELAAGEVGAQ 1680  
 QY 1681 SGFEKRDITPTEGRSTDAQGGKTSVITPELDNKAEEDGILAECLISAMPKGSHP 1740  
 Db 1681 SGFEKRDITPTEGRSTDAQGGKTSVITPELDNKAEEDGILAECLISAMPKGSHP 1740  
 QY 1741 FRVYKIMDQVQASASSAPNKQQLDGKKKCTSPVXPPIQNTERYTRVKNADSKNMLN 1800  
 Db 1741 FRVYKIMDQVQASASSAPNKQQLDGKKKCTSPVXPPIQNTERYTRVKNADSKNMLN 1800  
 QY 1801 AERVSDNKKQKQNLKANSKQPNDCLPNNDRVRSFAPDSPHHTPIEGTPIYCSRD 1860  
 Db 1801 AERVSDNKKQKQNLKANSKQPNDCLPNNDRVRSFAPDSPHHTPIEGTPIYCSRD 1860  
 QY 1861 SLSSLDPDDDDVLSREKALERAKENKESAKVTSHTLTIQNOQSAKTOAIAPQINR 1920  
 Db 1861 SLSSLDPDDDDVLSREKALERAKENKESAKVTSHTLTIQNOQSAKTOAIAPQINR 1920  
 QY 1921 GQPKPILQKOSTPQSSKDIIPRGAATDEKONPAIENTPVCFSHNSLSLSIDIDQEN 1980  
 Db 1921 GQPKPILQKOSTPQSSKDIIPRGAATDEKONPAIENTPVCFSHNSLSLSIDIDQEN 1980  
 QY 1981 NKENEPKETEPPDQSGEBSKQASGAPKSFHVEDPVCFSRNSLSLSIDSEDDLO 2040  
 Db 1981 NKENEPKETEPPDQSGEBSKQASGAPKSFHVEDPVCFSRNSLSLSIDSEDDLO 2040  
 QY 2041 ECISAMPKXKXSRILKGNKHSPPRMGILGEDTLTDLKDIORPDSHGLSPSENF 2100  
 Db 2041 ECISAMPKXKXSRILKGNKHSPPRMGILGEDTLTDLKDIORPDSHGLSPSENF 2100  
 QY 2101 WKALQGANISVSLHQAALAACTSRQASDSDSLISKSGISLQSPHILTDQEKPT 2160  
 Db 2101 WKALQGANISVSLHQAALAACTSRQASDSDSLISKSGISLQSPHILTDQEKPT 2160  
 QY 2161 SNKGPRILKPGKSTLETFKIIESESGIKGKKVYKSLITGVRNSSEISGQKQPLQAN 2220  
 Db 2161 SNKGPRILKPGKSTLETFKIIESESGIKGKKVYKSLITGVRNSSEISGQKQPLQAN 2220  
 QY 2221 MBSISGRMTLHPGVANSSSTSPVSKGKPLKTPAKSBSGOTATISPRGAPSVYS 2280  
 Db 2221 MBSISGRMTLHPGVANSSSTSPVSKGKPLKTPAKSBSGOTATISPRGAPSVYS 2280  
 QY 2281 ELSPVARQTSQIGSSKAPSRSGSRDSTPSRAQOPLSPPIOSPGRNSISPRNGISPPN 2340  
 Db 2281 ELSPVARQTSQIGSSKAPSRSGSRDSTPSRAQOPLSPPIOSPGRNSISPRNGISPPN 2340  
 QY 2341 KLSQLPRTSPSTASTKSSGSGKMYTSPGRQMSQNLTKQGLSKNASSIRSSASKG 2400  
 Db 2341 KLSQLPRTSPSTASTKSSGSGKMYTSPGRQMSQNLTKQGLSKNASSIRSSASKG 2400  
 QY 2401 LNMNNNGANKKVELSRWSTKSSGSSBDRBERPVLVQSFITKEABPPTLRKLBSBA 2460  
 Db 2401 LNMNNNGANKKVELSRWSTKSSGSSBDRBERPVLVQSFITKEABPPTLRKLBSBA 2460  
 QY 2461 SFESLSPSSRPAPSTRSQATPYLSPSLPDMSLJTHSSVQAGMRKLPNILEPTIEYNDG 2520

Db 2461 SFESISPSRSPASPTRSOAQTPVLSPLPDMSLSTHSSVQAGWRKLPNNISPTIEYNDG 2520  
QY 2521 RPAKHDIARSHSESPRLPINRSGTWKREHSKHSSSLPRYSTWRRTGSSSITLSASSES 2580  
Db 2521 RPAKHDIARSHSESPRLPINRSGTWKREHSKHSSSLPRYSTWRRTGSSSITLSASSES 2580  
QY 2581 SEKAKSEDEKHVNSISGTQSKENQVSAKGTWRKIKENEFSPTNSTSGTVSSGATNGAES 2640  
Db 2581 SEKAKSEDEKHVNSISGTQSKENQVSAKGTWRKIKENEFSPTNSTSGTVSSGATNGAES 2640  
QY 2641 KTLIYQMAPAVSKTEDVWVRTEDCPINNPRSGRSPGTGTPVIDSVSEKANPNIKDSKDN 2700  
Db 2641 KTLIYQMAPAVSKTEDVWVRTEDCPINNPRSGRSPGTGTPVIDSVSEKANPNIKDSKDN 2700  
QY 2701 QAKQVNGSVPMRTVGLNRLNSFIQVADAPDQKTEIKPGQNNPVPVSEITNESSIYERT 2760  
Db 2701 QAKQVNGSVPMRTVGLNRLNSFIQVADAPDQKTEIKPGQNNPVPVSEITNESSIYERT 2760  
QY 2761 PFSSSSSSSKHSSPSGTVAAARVTPNNYNPSPRKSADSTASAPSOIPTPVNNNTKKRDSKT 2820  
Db 2761 PFSSSSSSSKHSSPSGTVAAARVTPNNYNPSPRKSADSTASAPSOIPTPVNNNTKKRDSKT 2820  
QY 2821 DSTESSGTQSPKRHSGSYLVTSTV 2843  
Db 2821 DSTESSGTQSPKRHSGSYLVTSTV 2843

Search completed: August 25, 2004, 17:25:57  
Job time : 52.5 secs

**This Page Blank (uspio)**